



Faculté
de Médecine

Aix-Marseille Université



Sciences Economiques et Sociales de la
Santé & Traitement de l'Information Médicale

Inserm / IRD / Aix-Marseille Université

Tests Paramétriques

Plan

1. Rappels: principe des tests
2. 2 moyennes observées
3. 2 pourcentages observés
4. Corrélation
5. Plusieurs moyennes observées
6. Plusieurs pourcentages observés

Mode

1. Théorie
2. Exercice Dirigé, *logiciel R*
3. Exercice Individuel, *logiciel R*

I. Rappels

1. Rappels

2. 2 Moyennes

3. 2 Pourcent.

4. Corrélation

5. x Moyennes

6. x Pourcent.

■ Variabilité individuelle:

⇒ Individus tous différents

⇒ Variable: mesures différentes /individus

⇒ Distribution de la variable

1. Rappels

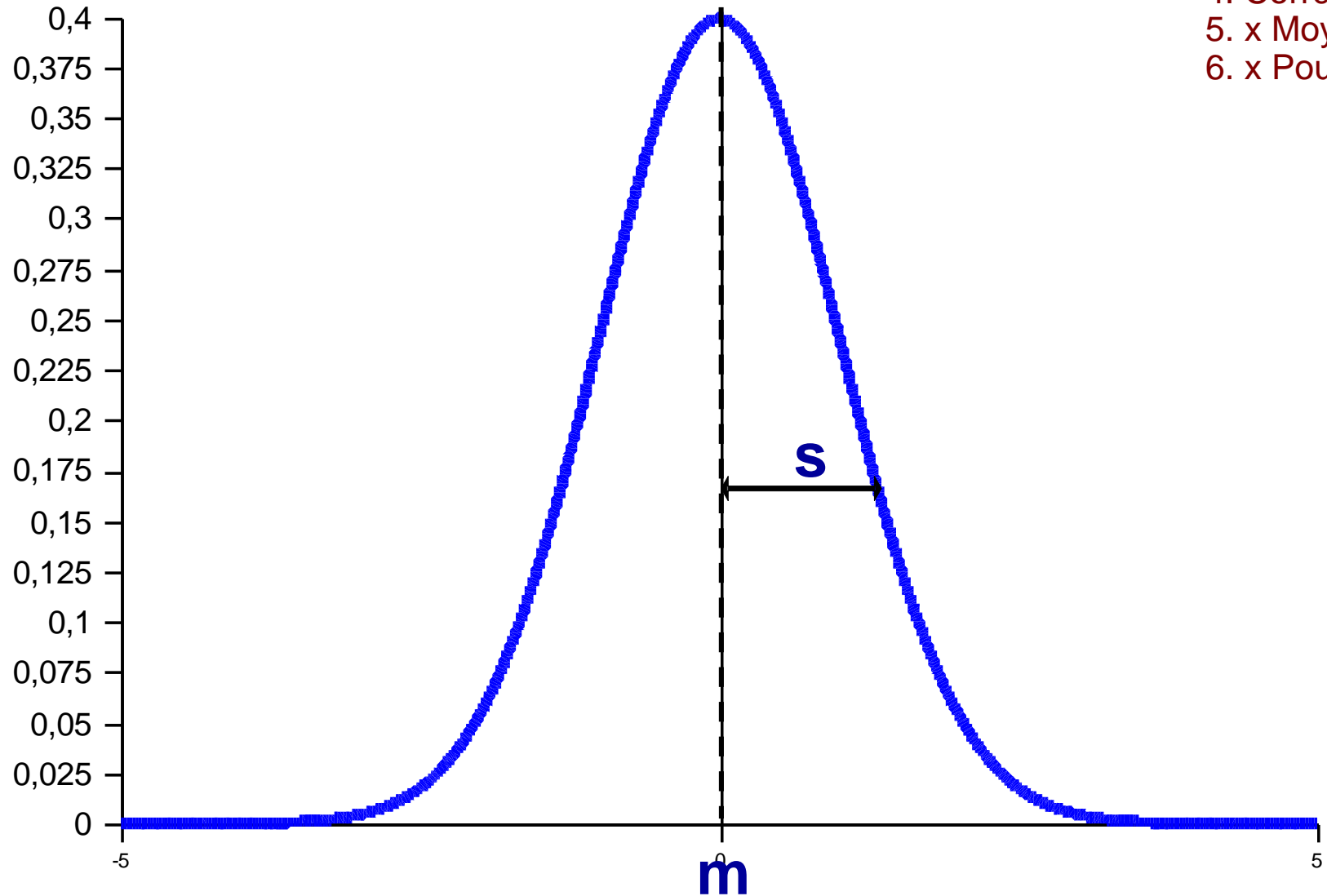
2. 2 Moyennes

3. 2 Pourcent.

4. Corrélation

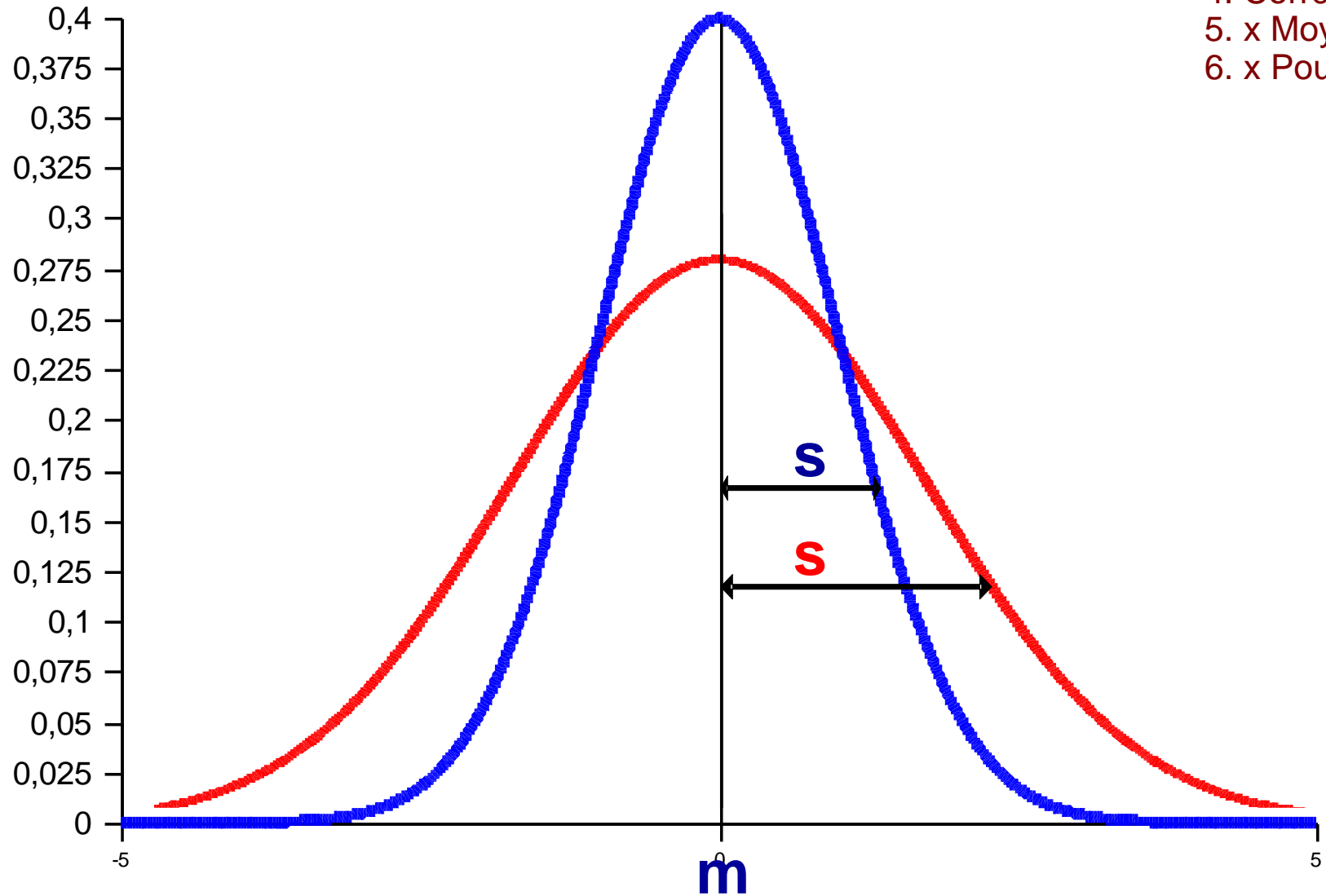
5. x Moyennes

6. x Pourcent.



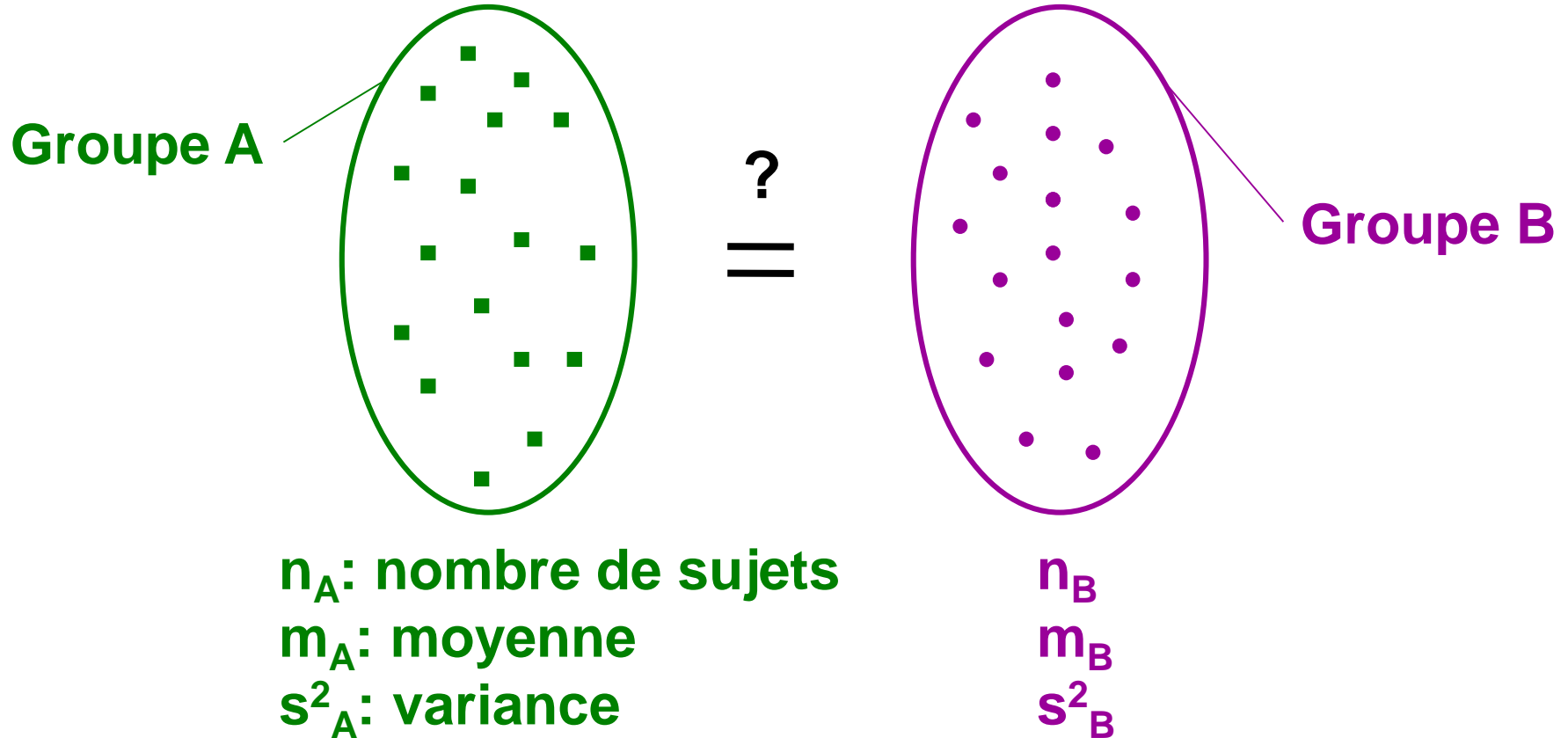
1. Rappels

2. 2 Moyennes
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.



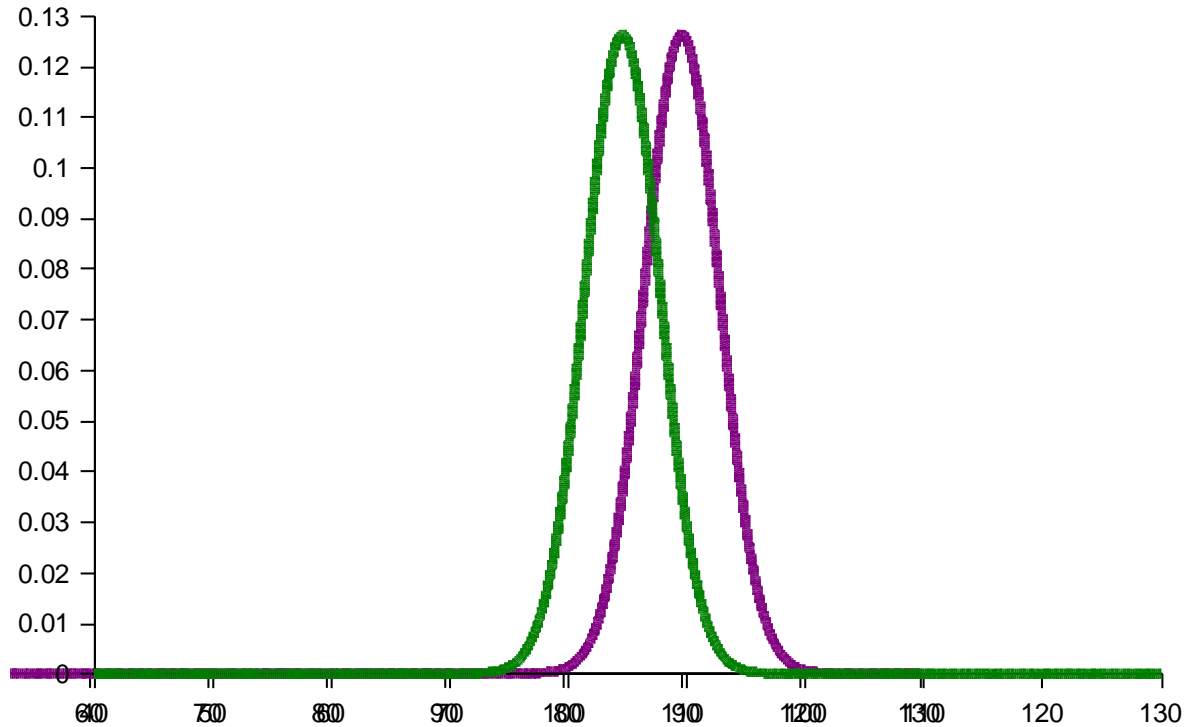
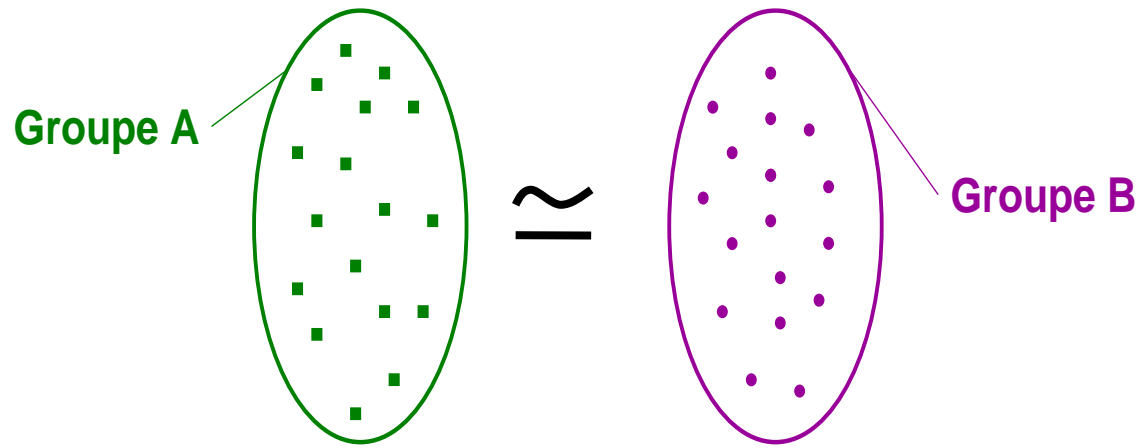
1. Rappels

2. 2 Moyennes
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.



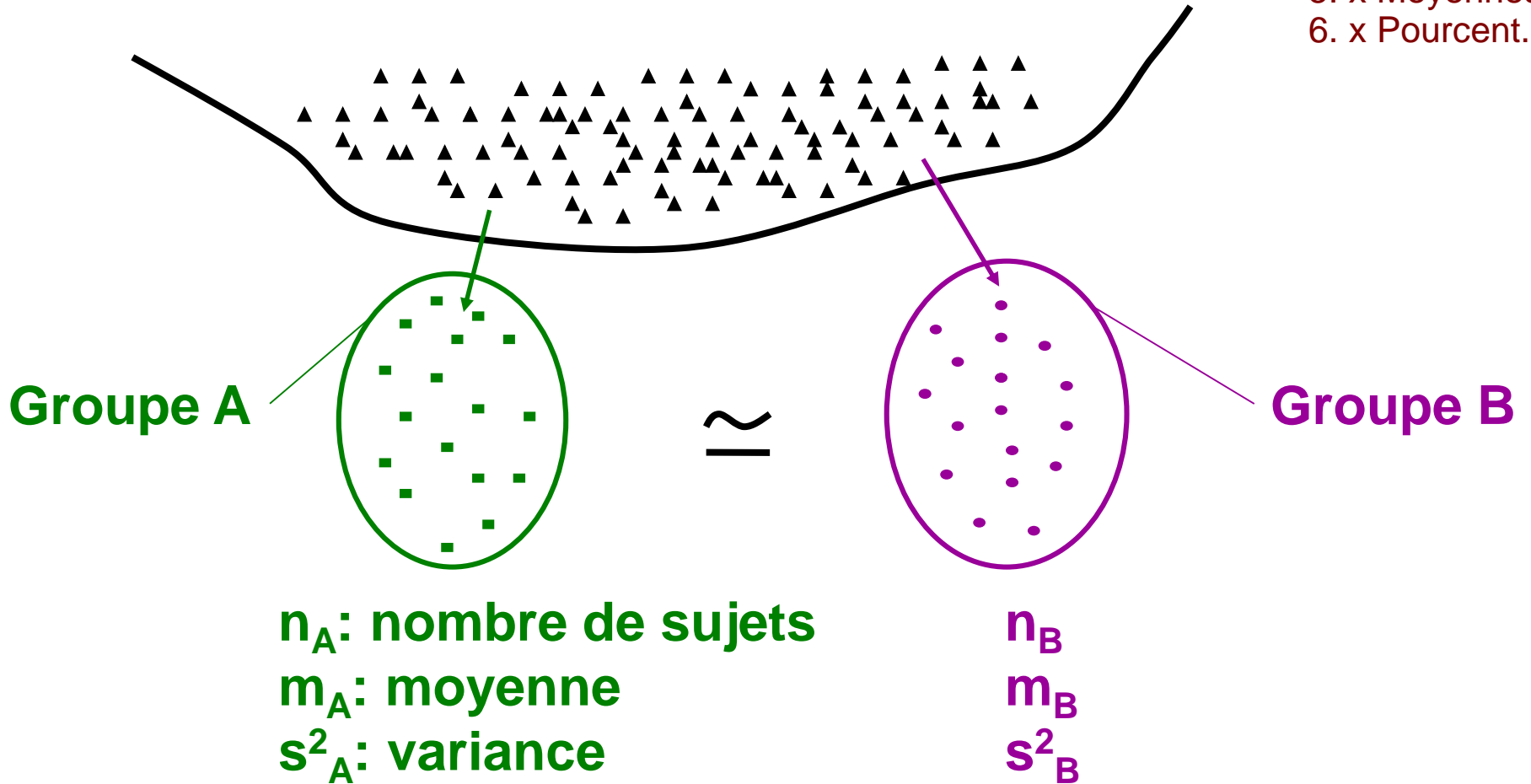
1. Rappels

2. 2 Moyennes
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.



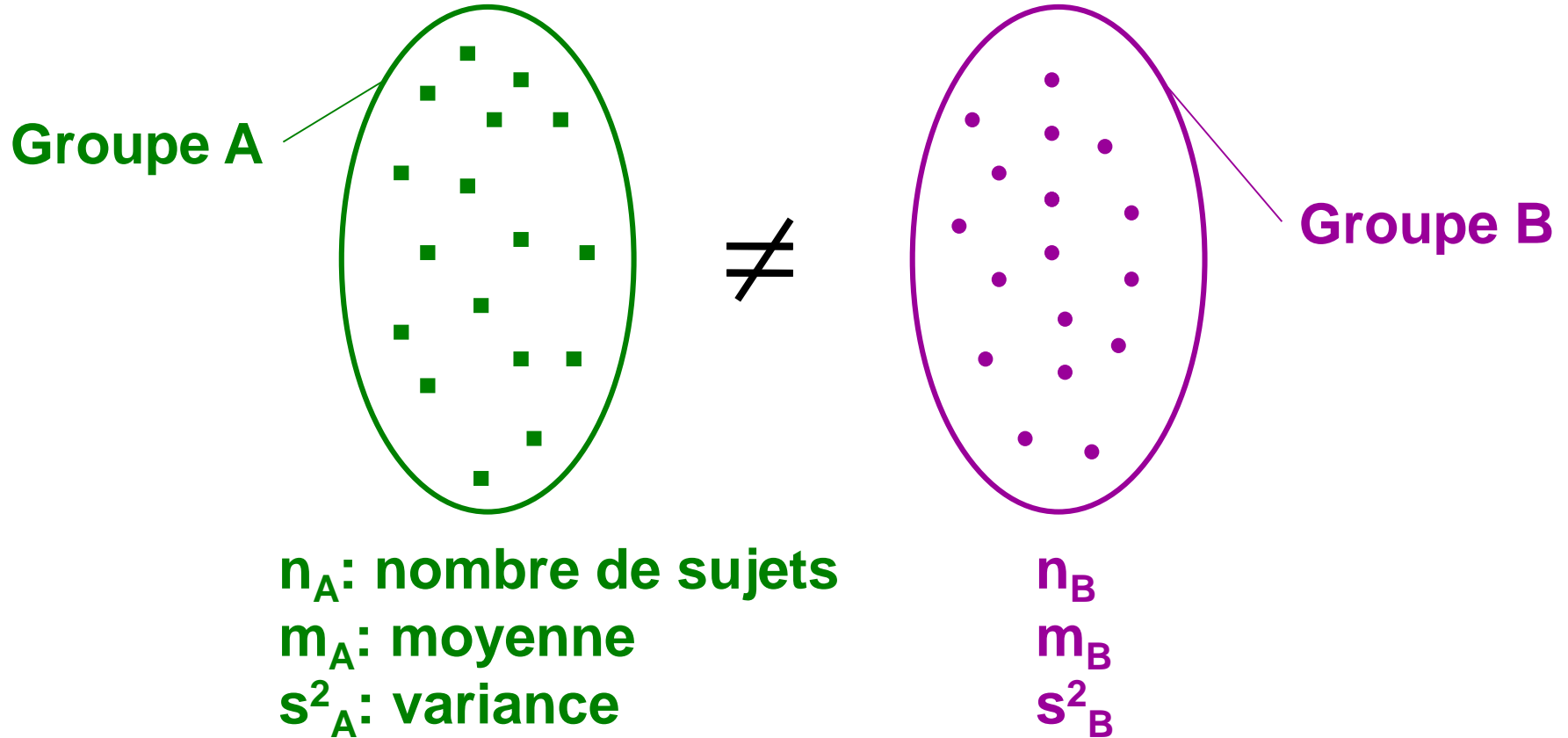
1. Rappels

2. 2 Moyennes
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.



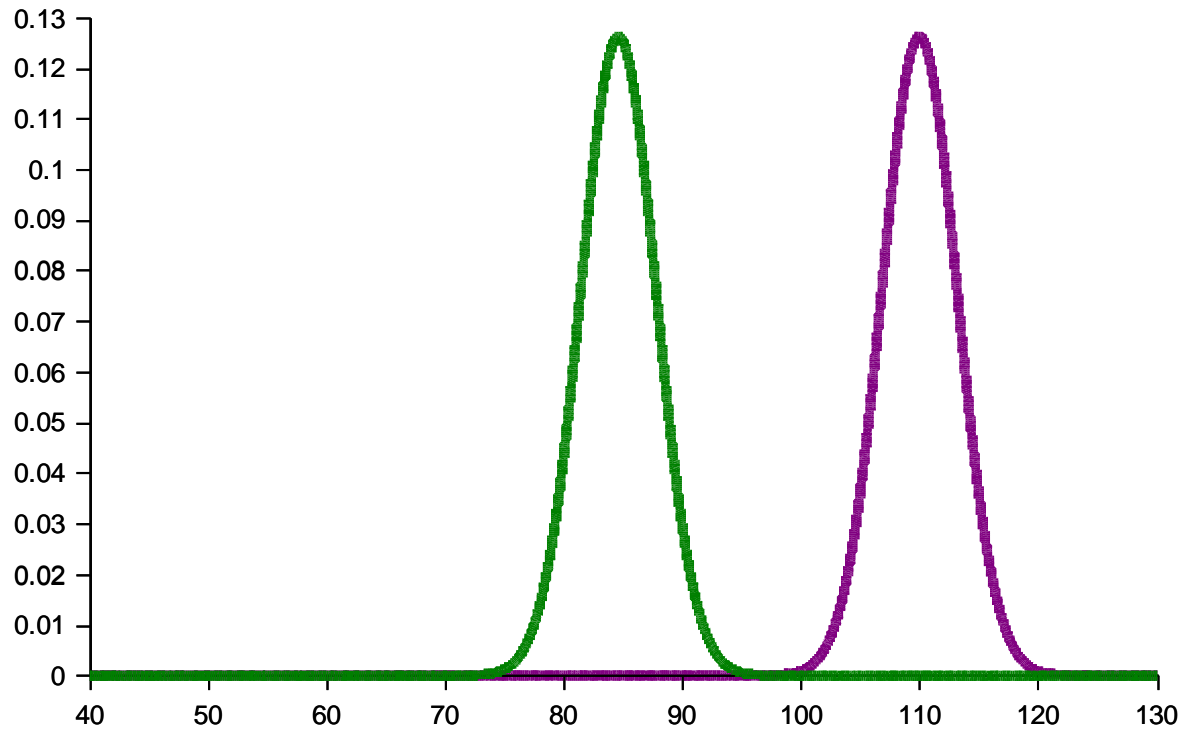
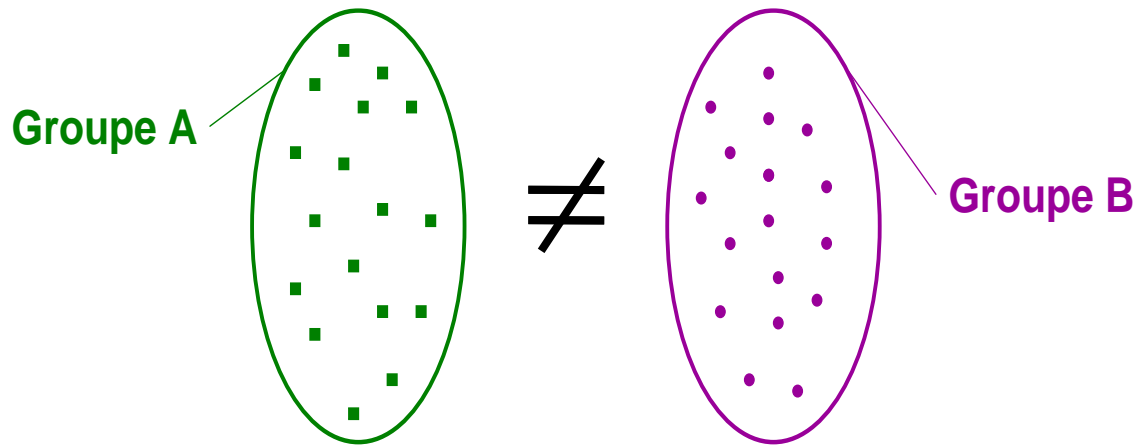
1. Rappels

2. 2 Moyennes
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.



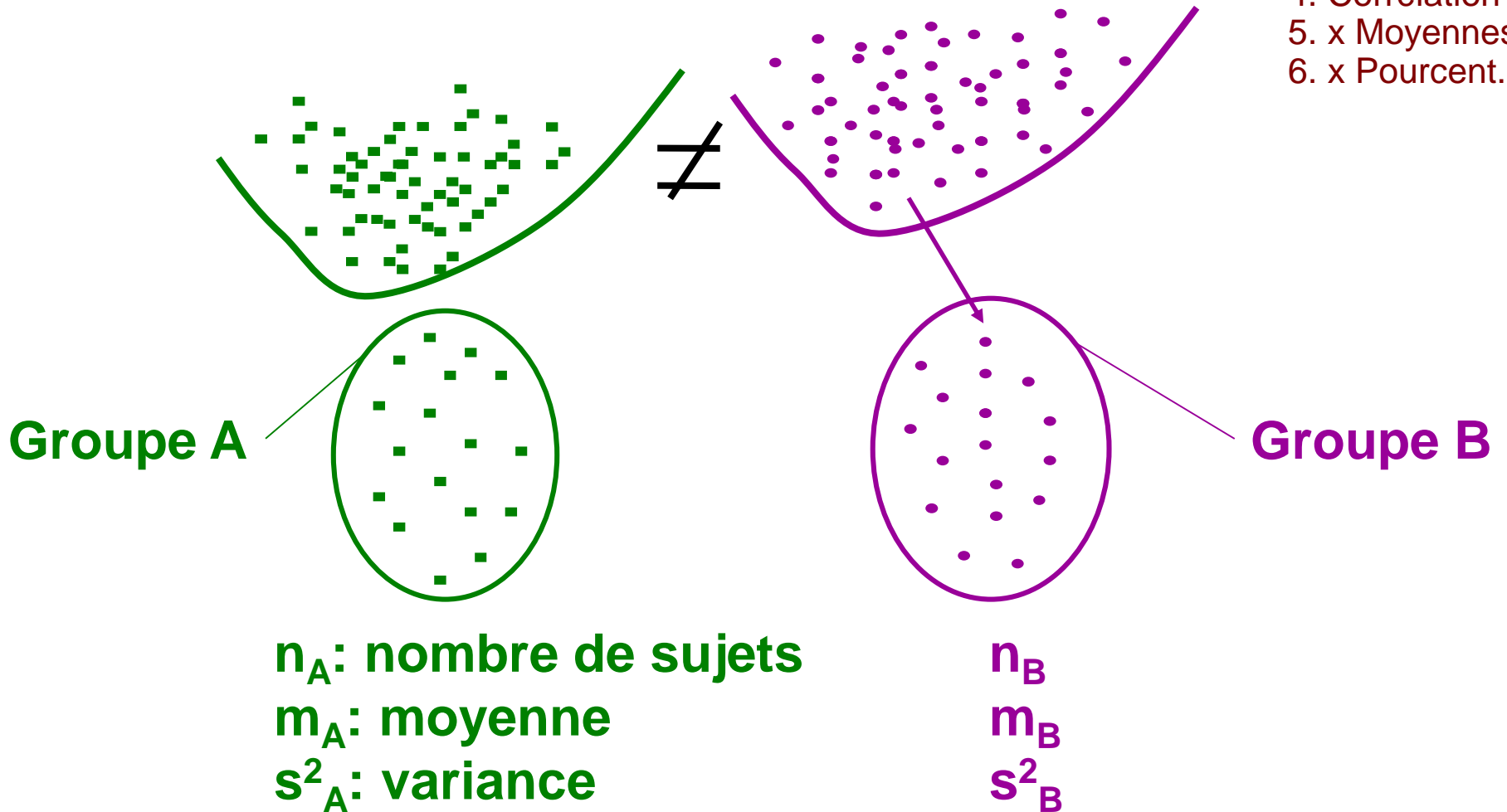
1. Rappels

- 2. 2 Moyennes
- 3. 2 Pourcent.
- 4. Corrélation
- 5. x Moyennes
- 6. x Pourcent.



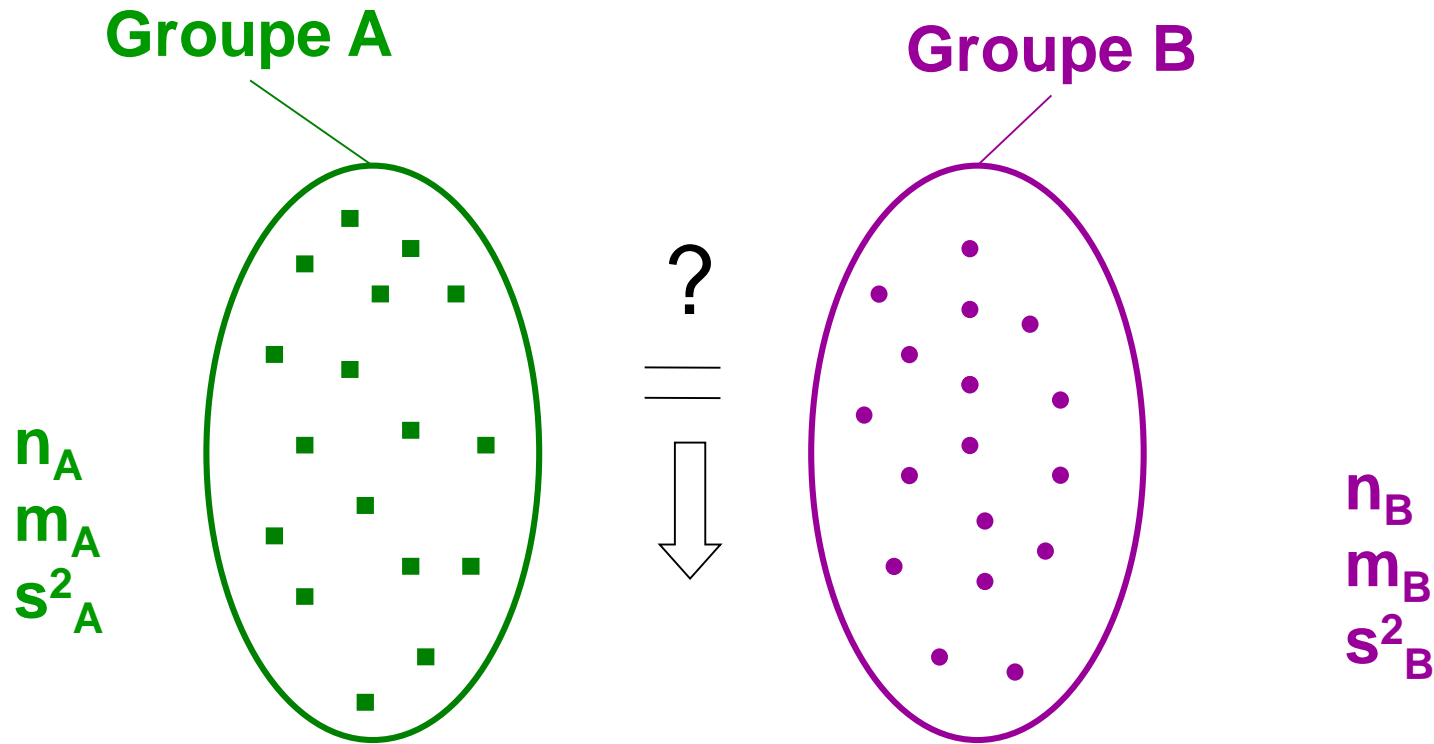
1. Rappels

2. 2 Moyennes
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.



1. Rappels

2. 2 Moyennes
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.



Comparer le critère de jugement
mesuré sur A
mesuré sur B

II. Comparaison de 2 moyennes observées

1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

Exemple: Effet du Tabac sur la tension artérielle,
Echantillon de 32 sujets

- Importer le fichier de données *TABAC.xls*
- Moyenne globale de la TAS
 $m = ?$ mmHg
- Variance globale de la TAS
 $s^2 = ?$ mmHg²
- Graphiques

II. Comparaison de 2 moyennes observées

1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

Exemple: Effet du Tabac, Echantillon de 32 sujets

- Importer le fichier de données *TABAC.xls*
data<-read.csv2("D:\\BIOSTAT\\TABAC.csv", header=TRUE)
- Moyenne globale de la TAS
m= ? mmHg *mean(data\$TAS)*
- Variance globale de la TAS
s²= ? mmHg² *var(data\$TAS)*
- Graphiques *hist(data\$TAS)* *boxplot(data\$TAS)*

II. Comparaison de 2 moyennes observées

1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

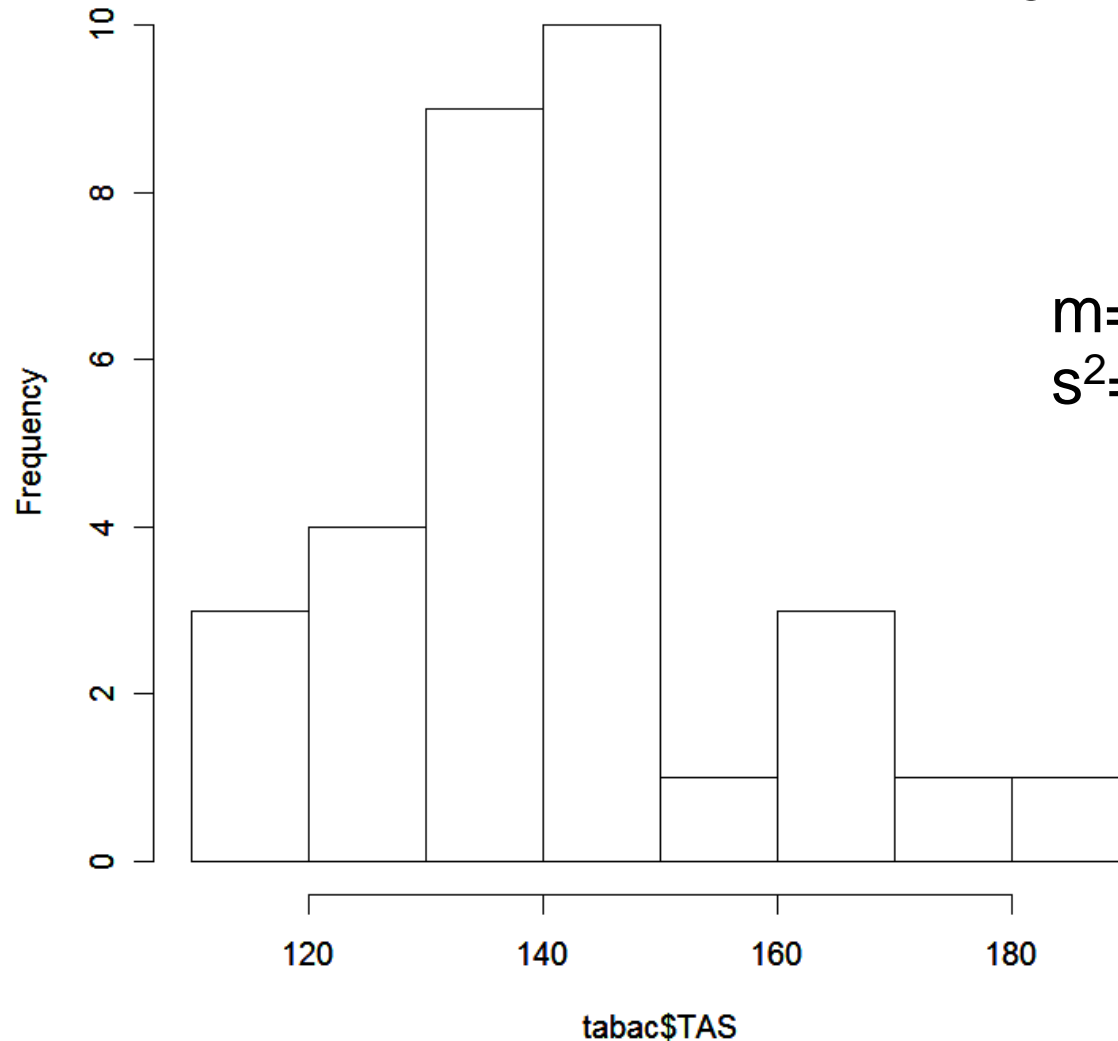
Exemple: Effet du Tabac, Echantillon de 32 sujets

- Importer le fichier de données *TABAC.xls*
data<-read.csv2("D:\\BIOSTAT\\TABAC.csv", header=TRUE)
- Moyenne globale de la TAS
 $m = 140,8$ mmHg *mean(data\$TAS)*
- Variance globale de la TAS
 $s^2 = 252,9$ mmHg² *var(data\$TAS)*
- Graphiques *hist(data\$TAS)* *boxplot(data\$TAS)*

hist(tabac\$TAS),

Histogram of tabac\$TAS

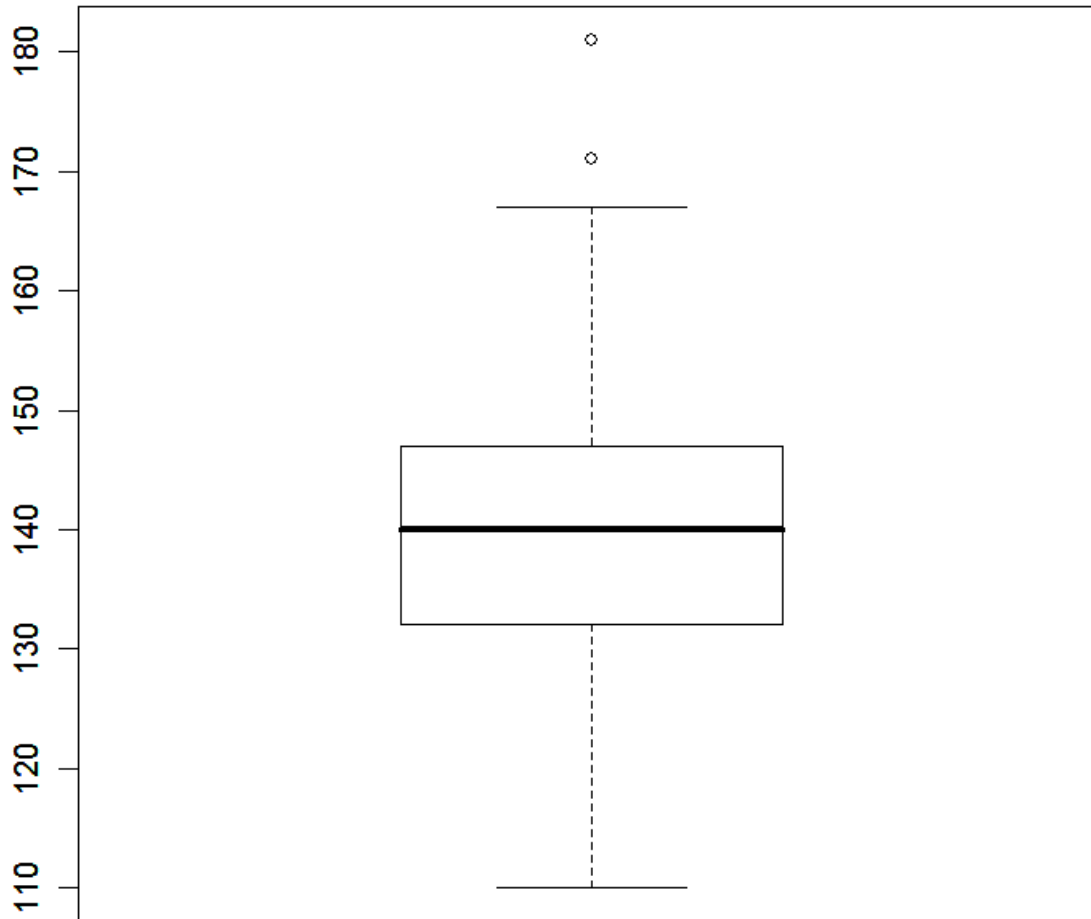
N=32



$m=140,8$ mmHg
 $s^2=252,9$

1. Rappels
2. 2 Moyennes
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

boxplot(tabac\$TAS),



N=32

$m=140,8$ mmHg
 $s^2=252,9$

1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélacion
5. x Moyennes
6. x Pourcent.

Effet du Tabac, Echantillon de 32 sujets

■ 2 groupes

17 fumeurs $m_f = ?$ mmHg
 $s_f^2 = ?$ mmHg²
graphiques

attach(data)

mean(TAS[Tabac==1])
var(TAS[Tabac==1])

hist(TAS[Tabac==1], col='red')

15 non fumeurs

$m_{nf} = ?$ mmHg
 $s_{nf}^2 = ?$ mmHg²
graphiques

mean(TAS[Tabac==0])

var(TAS[Tabac==0])

hist(TAS[Tabac==0], col='blue')

boxplot(TAS~Tabac)

Effet du Tabac, Echantillon de 32 sujets

■ 2 groupes

17 fumeurs $m_f = 149,6$ mmHg
 $s_f^2 = 212,1$ mmHg²
graphiques

attach(data)

mean(TAS[Tabac==1])
var(TAS[Tabac==1])

hist(TAS[Tabac==1], col='red')

15 non fumeurs

$m_{nf} = 130,9$ mmHg
 $s_{nf}^2 = 118,1$ mmHg²
graphiques

mean(TAS[Tabac==0])

var(TAS[Tabac==0])

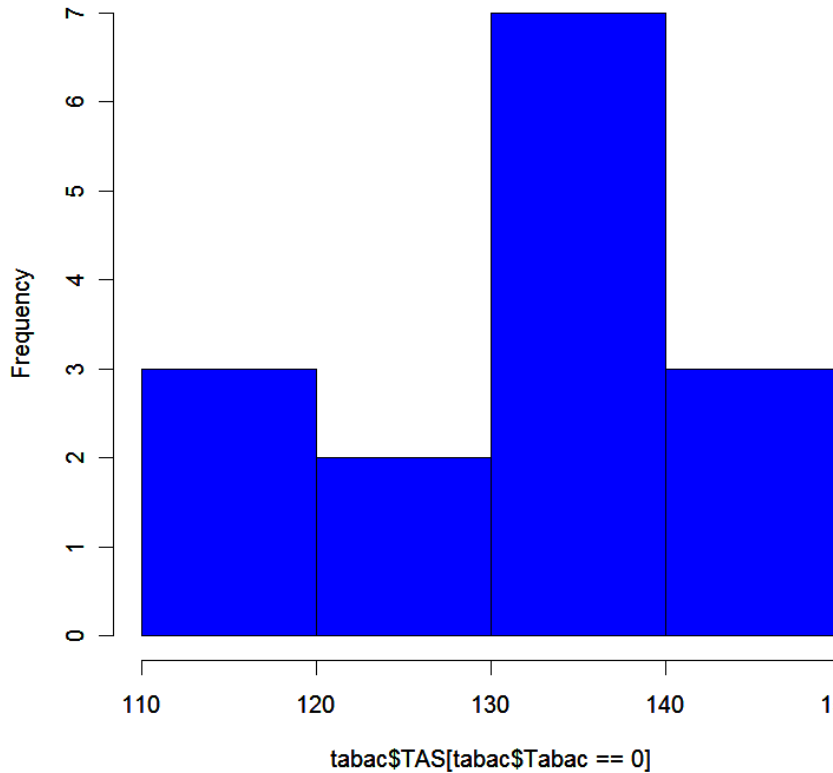
hist(TAS[Tabac==0], col='blue')

boxplot(TAS~Tabac)

1. Rappels
2. 2 Moyennes
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

15 non fumeurs

Histogram of `tabac$TAS[tabac$Tabac == 0]`

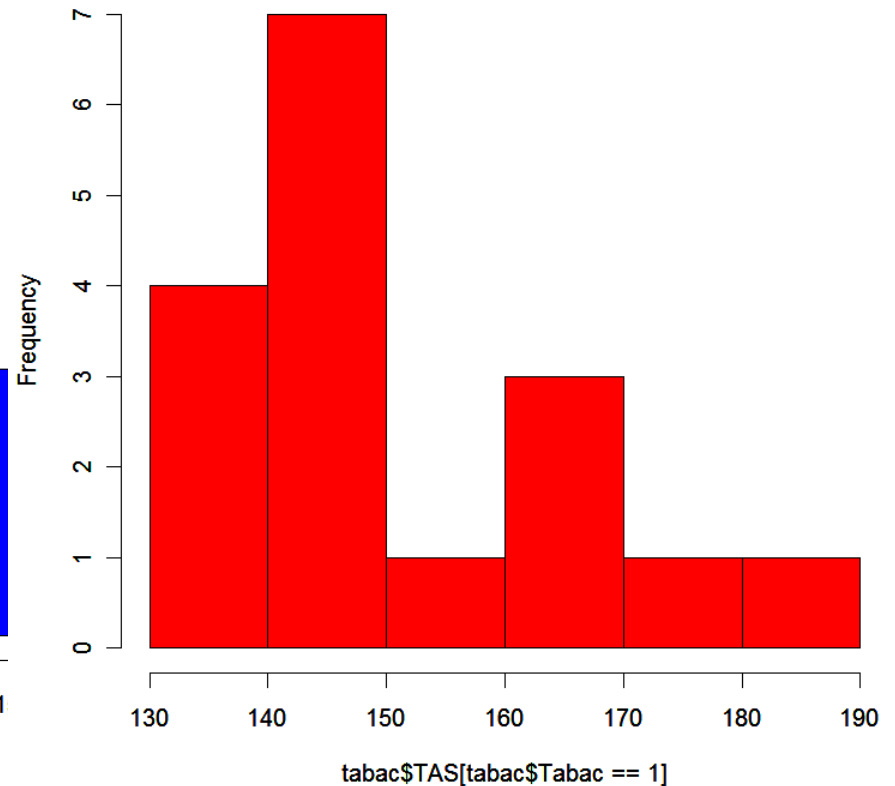


$$m_{nf} = 130,9 \text{ mmHg}$$

$$s_f^2 = 118,1$$

17 fumeurs

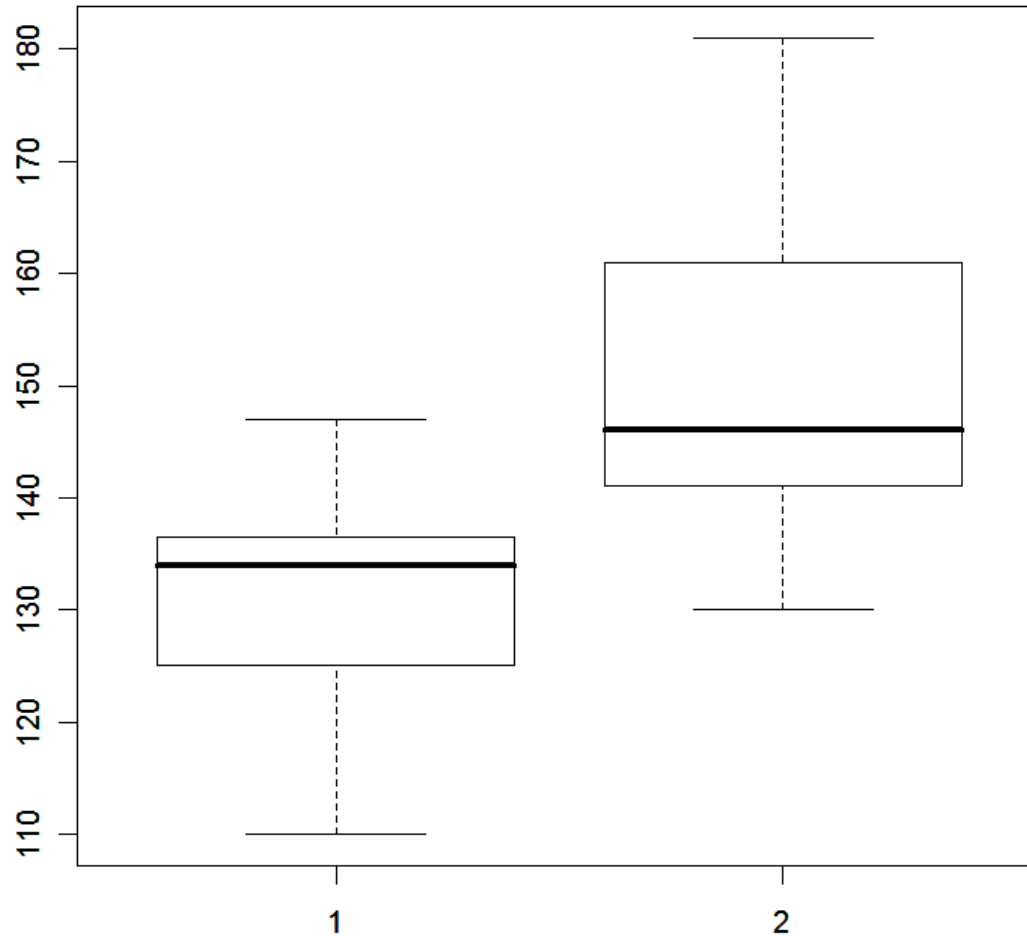
Histogram of `tabac$TAS[tabac$Tabac == 1]`



$$m_f = 149,6 \text{ mmHg}$$

$$s_f^2 = 212,1$$

1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.



$m_{nf}=130,9$ mmHg
 $s_f^2=118,1$

$m_f=149,6$ mmHg
 $s_f^2=212,1$

1. Hypothèses:

H0: $\mu_{NF} = \mu_F$ la TAS est en moyenne identique chez les fumeurs et les non fumeurs

H1: $\mu_{NF} \neq \mu_F$ la TAS moyenne est différente chez les fumeurs et les non fumeurs

2. Prédiction sous H0:

Sous H0 et si les conditions d'applications sont respectées, alors

$$T = \frac{|\mu_{NF} - \mu_F|}{\sqrt{\sigma^2 \left(\frac{1}{n_{NF}} + \frac{1}{n_F} \right)}} \rightarrow T_{n_{NF} + n_F - 2}$$

Loi de Student

1. Hypothèses
2. Prédiction sous H_0 :

Conditions d'applications

→ $X \rightarrow N(\mu; \sigma^2)$ **OU** $n_{NF} > 30$ **ET** $n_F > 30$

ET

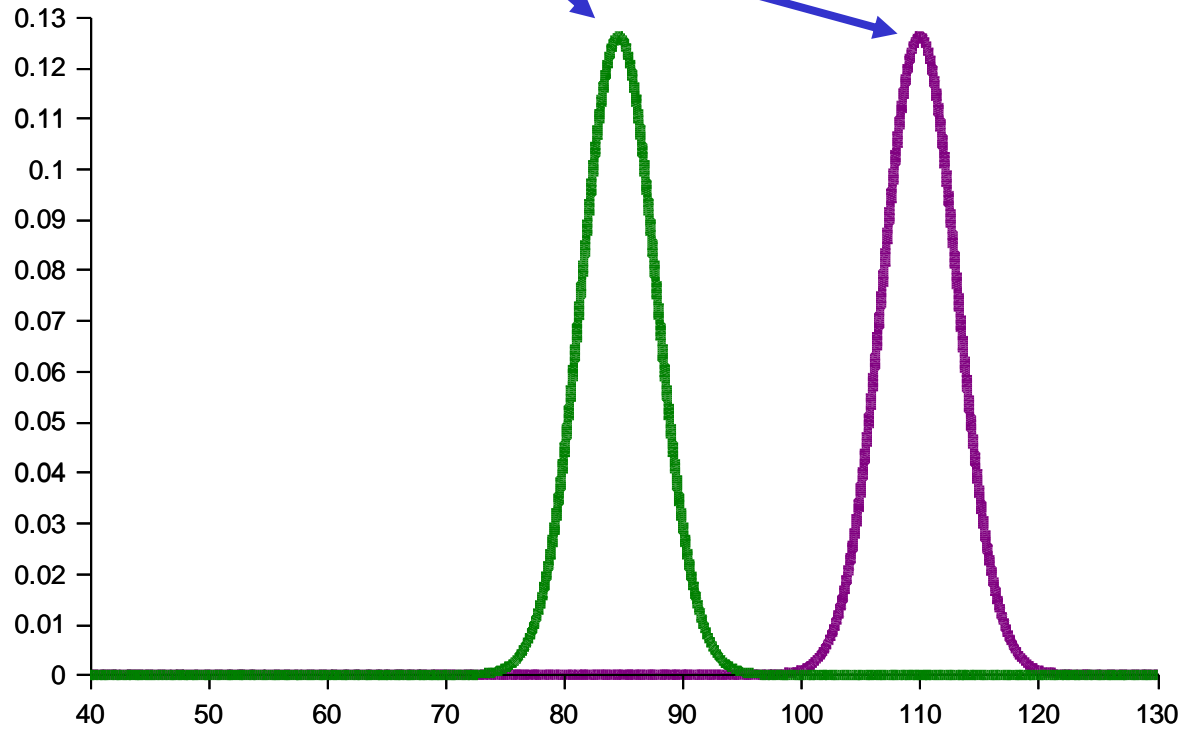
→ $\sigma_{NF}^2 = \sigma_F^2$

ET

→ Indépendance des individus

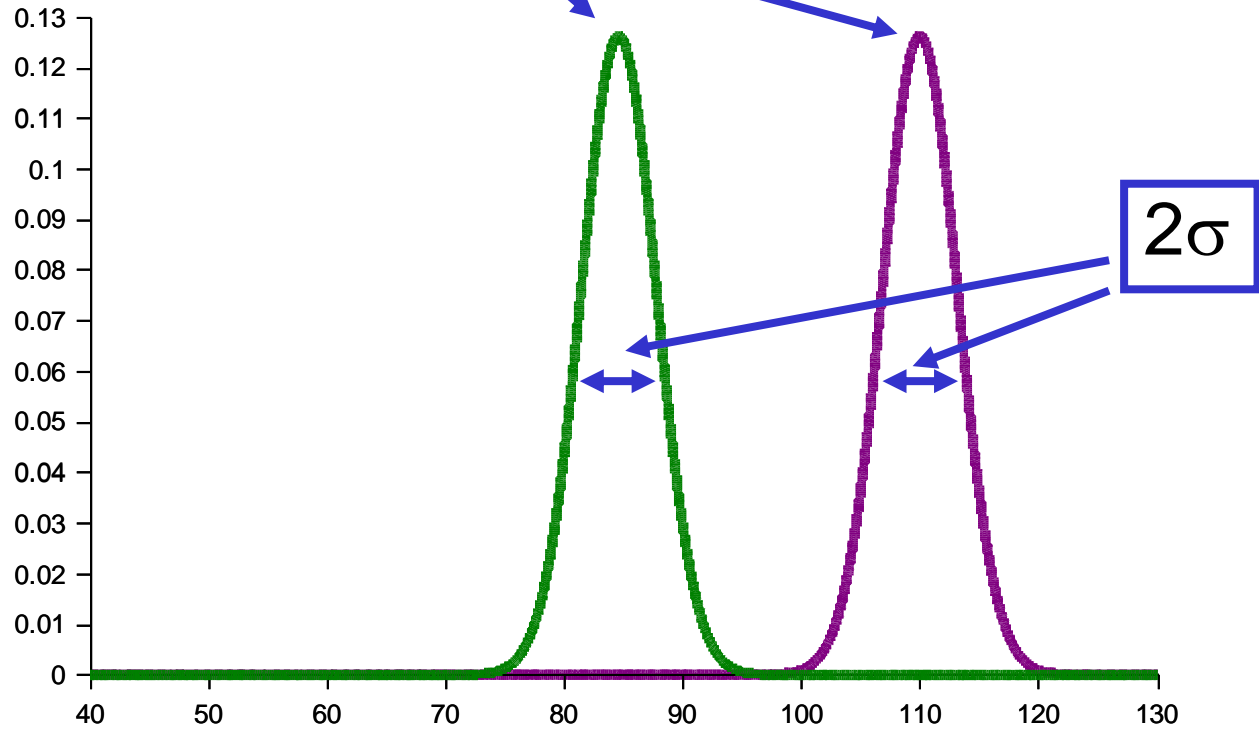
1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

Loi Normale



1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

Loi Normale



1. Hypothèses
2. Prédiction sous H0
3. Confrontation

Valeurs observées = estimations

$$t_o = \frac{|m_{NF} - m_F|}{\sqrt{s^2 \left(\frac{1}{n_{NF}} + \frac{1}{n_F} \right)}} \leftrightarrow T_{\alpha, \nu}$$

degrés de liberté (ddl)

seuil pour un risque donné

variance commune estimée
 à partir des variances estimées sur chaque échantillon

$$s^2 = \frac{(n_{NF} - 1)s_{NF}^2 + (n_F - 1)s_F^2}{n_{NF} + n_F - 2}$$

1. Hypothèses
2. Prédiction sous H0
3. Confrontation **test de Student**

t.test(TAS[Tabac==0], TAS[Tabac==1], var.equal=TRUE)

1. Hypothèses
2. Prédiction sous H0
3. Confrontation **test de Student**

t.test(TAS[Tabac==0], TAS[Tabac==1], var.equal=TRUE)

Test de Student

1. Hypothèses
2. Prédiction sous H0
3. Confrontation **test de Student**

```
t.test(TAS[Tabac==0], TAS[Tabac==1], var.equal=TRUE)
```



Tension artérielle chez les non fumeurs

1. Hypothèses
2. Prédiction sous H0
3. Confrontation **test de Student**

```
t.test(TAS[Tabac==0], TAS[Tabac==1], var.equal=TRUE)
```

Tension artérielle chez les fumeurs

1. Hypothèses
2. Prédiction sous H0
3. Confrontation **test de Student**

```
t.test(TAS[Tabac==0], TAS[Tabac==1], var.equal=TRUE)
```

Condition d'égalité des variance

1. Hypothèses
2. Prédiction sous H0
3. Confrontation

t.test(TAS[Tabac==0], TAS[Tabac==1], var.equal=TRUE)

Two Sample t-test

```
data: TAS[Tabac == 0] and TAS[Tabac == 1]
t = -4.0742, df = 30, p-value = 0.0003113
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -28.106070 -9.337067
sample estimates:
mean of x mean of y
130.8667 149.5882
```


1. Hypothèses
2. Prédiction sous H0
3. Confrontation

```
t.test(TAS[Tabac==0],TAS[Tabac==1], var.equal=TRUE)
```

Two Sample t-test

Test de Student pour 2 échantillons indépendants

```
data: TAS[Tabac == 0] and TAS[Tabac == 1]
t = -4.0742, df = 30, p-value = 0.0003113
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-28.106070 -9.337067
sample estimates:
mean of x mean of y
130.8667 149.5882
```

1. Hypothèses
2. Prédiction sous H0
3. Confrontation

```
t.test(TAS[Tabac==0], TAS[Tabac==1], var.equal=TRUE)
```

Two Sample t-test

data: TAS[Tabac == 0] and TAS[Tabac == 1]

t = -4.0742, df = 30, p-value = 0.0003113

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

95 percent confidence interval:

-28.106070 -9.337067

sample estimates:

mean of x mean of y

130.8667 149.5882

données

1. Hypothèses
2. Prédiction sous H0
3. Confrontation

```
t.test(TAS[Tabac==0], TAS[Tabac==1], var.equal=TRUE)
```

Two Sample t-test

```
data: TAS[Tabac == 0] and TAS[Tabac == 1]
t = -4.0742, df = 30, p-value = 0.000441
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -28.106070 -9.337067
sample estimates:
mean of x mean of y
130.8667 149.5882
```

t₀ calculé sous H0

1. Hypothèses
2. Prédiction sous H0
3. Confrontation

```
t.test(TAS[Tabac==0], TAS[Tabac==1], var.equal=TRUE)
```

Two Sample t-test

```
data: TAS[Tabac == 0] and TAS[Tabac == 1]
t = -4.0742, df = 30, p-value = 0.0003113
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-28.106070 -9.337067
sample estimates:
mean of x mean of y
130.8667 149.5882
```

Petit « p »

1. Hypothèses
2. Prédiction sous H0
3. Confrontation

t.test(TAS[Tabac==0], TAS[Tabac==1], var.equal=TRUE)

Two Sample t-test

data: TAS[Tabac == 0] and TAS[Tabac == 1]

t = -4.0742, df = 30, p-value = 0.0003113

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

95 percent confidence interval:

-28.106070 -9.337067

sample estimates:

mean of x mean of y

130.8667 149.5882

H1

1. Hypothèses
2. Prédiction sous H0
3. Confrontation
4. Interprétation

- ➔ $p < 0,05$
- ➔ Test significatif
- ➔ On rejette H0, au risque $\alpha = 5\%$
- ➔ Il y a une différence entre les 2 groupes
- ➔ Dans le sens **“les fumeurs ont une TAS moyenne plus élevée que les non-fumeurs”**

Vérification des Conditions d'application

■ Normalité

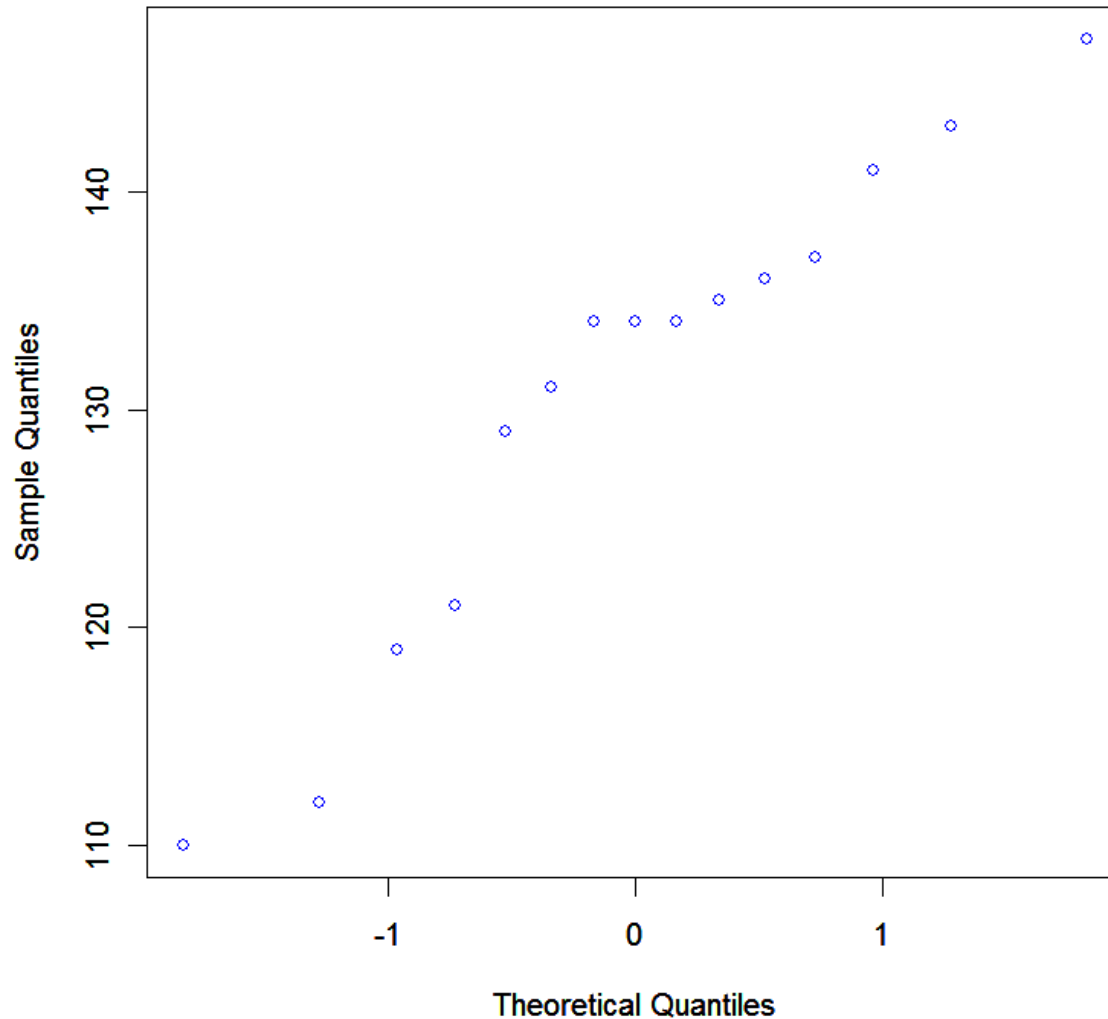
- Histogrammes

-  □ qq-plot: `qqnorm(TAS [Tabac==0], col="blue")`

- Test non paramétrique

```
qqnorm(TAS [Tabac==0], col="blue")
```

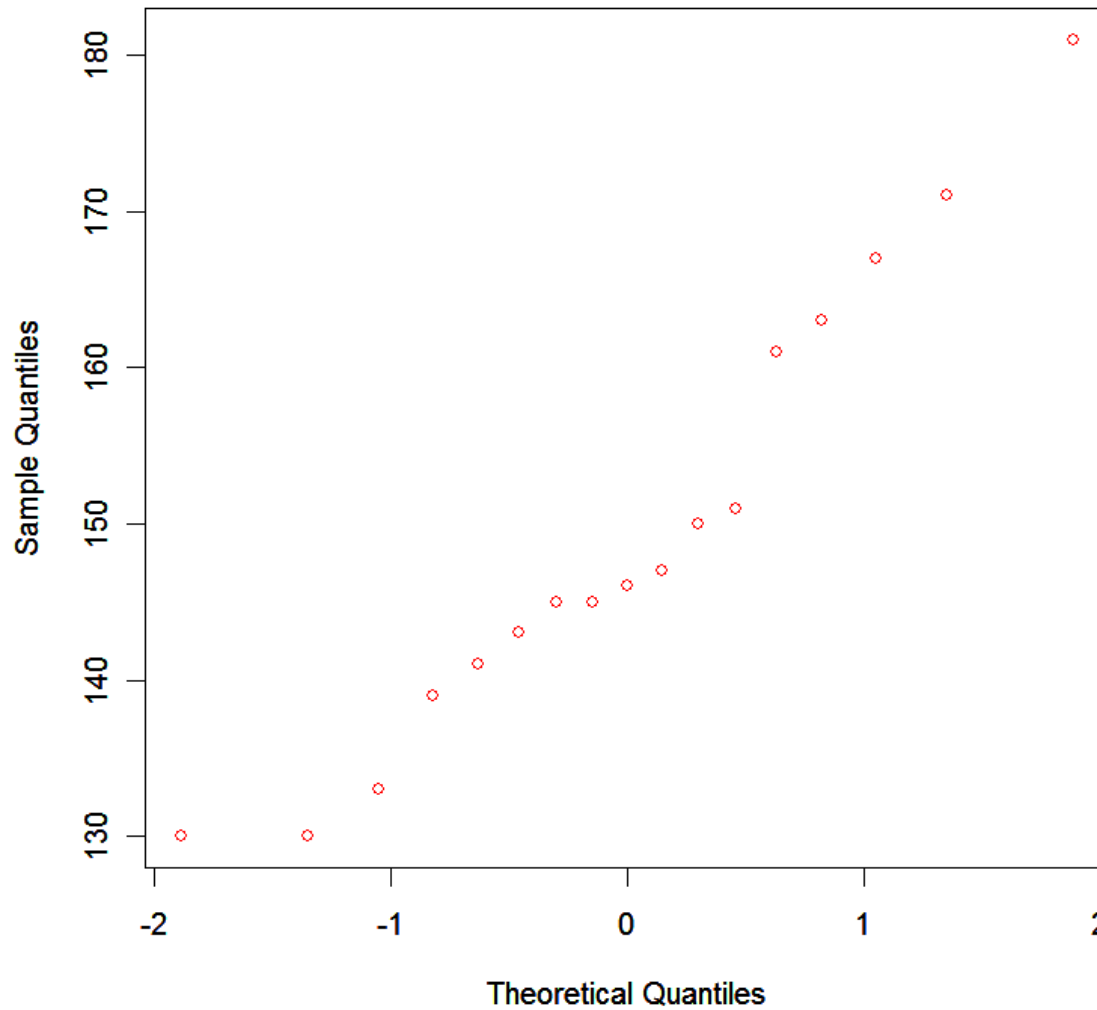
Normal Q-Q Plot



1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

qqnorm(TAS [Tabac==1], col="red")

Normal Q-Q Plot



- 1. Rappels
- 2. 2 Moyennes**
- 3. 2 Pourcent.
- 4. Corrélation
- 5. x Moyennes
- 6. x Pourcent.

Vérification des Conditions d'application

■ Normalité

- Histogrammes
- qq-plot: *qqnorm(TAS [Tabac==0], col="blue")*
- Test non paramétrique

■ Egalité des variances

- Test de comparaison de 2 variances
 - H0: égalité
 - H1: différence

var.test(TAS[Tabac==0], TAS[Tabac==1])

ou var.test(TAS~Tabac)

■ Indépendance

```
var.test(TAS[Tabac==0], TAS[Tabac==1])
```

F test to compare two variances

```
data: TAS[Tabac == 0] and TAS[Tabac == 1]
F = 0.5568, num df = 14, denom df = 16, p-value = 0.2773
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1976701 1.6278629
sample estimates:
ratio of variances
 0.5568401
```

1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélacion
5. x Moyennes

```
var.test(TAS[Tabac==0], TAS[Tabac==1])
```

F test to compare two variances

Comparaison de 2 variances

```
data: TAS[Tabac == 0] and TAS[Tabac == 1]
F = 0.5568, num df = 14, denom df = 16, p-value = 0.2773
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1976701 1.6278629
sample estimates:
ratio of variances
 0.5568401
```

1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélacion
5. x Moyennes
6. x Pourcent.

```
var.test(TAS[Tabac==0], TAS[Tabac==1])
```

F test to compare two variances

data: TAS[Tabac == 0] and TAS[Tabac == 1]
F = 0.5568, num df = 14, denom df = 16, p-value = 0.2773
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
0.1976701 1.6278629
sample estimates:
ratio of variances
0.5568401

Données

1. Rappels
2. 2 Moyennes
3. 2 Pourcent.
4. Corrélacion
5. x Moyennes
6. x Pourcent.

var.test(TAS[Tabac==0], TAS[Tabac==1])

F test to compare two variances

data: TAS[Tabac == 0] and TAS[Tabac == 1]
 F = 0.5568, num df = 14, denom df = 14, p-value = 0.2773
 alternative hypothesis: true ratio of variances is not equal to 1
 95 percent confidence interval:
 0.1976701 1.6278629
 sample estimates:
 ratio of variances
 0.5568401

F₀ calculé sous H₀

1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélacion
5. x Moyennes
6. x Pourcent.

```
var.test(TAS[Tabac==0], TAS[Tabac==1])
```

F test to compare two variances

```
data: TAS[Tabac == 0] and TAS[Tabac == 1]
F = 0.5568, num df = 14, denom df = 16, p-value = 0.2773
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1976701 1.6278629
sample estimates:
ratio of variances
 0.5568401
```

Petit « p »



1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

```
var.test(TAS[Tabac==0], TAS[Tabac==1])
```

F test to compare two variances

data: TAS[Tabac == 0] and TAS[Tabac == 1]

~~F = 0.5568, num df = 14, denom df = 16, p-value = 0.2773~~

~~alternative hypothesis: true ratio of variances is not equal to 1~~

~~95 percent confidence interval.~~

~~0.1976701 1.6278629~~

sample estimates:

ratio of variances

0.5568401

H1




```
var.test(TAS[Tabac==0], TAS[Tabac==1])
```

F test to compare two variances

```
data: TAS[Tabac == 0] and TAS[Tabac == 1]
F = 0.5568, num df = 14, denom df = 16, p-value = 0.2773
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1976701 1.6278629
sample estimates:
ratio of variances
 0.5568401
```

- ➔ $p > 0,05$
- ➔ Test Non significatif
- ➔ Non rejet de H_0 au risque β
- ➔ On ne met pas en évidence de différence significative entre les 2 variances

Exercice

1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

- fichier *TABAC.csv*
- Y a-t-il une différence entre la moyenne de la TAS des hommes (code 1) et celle des femmes (code 0) ?

Exercice

■ Statistiques descriptives

1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

```
data<-read.csv2("D:\\BIOSTAT\\TABAC.csv", header=TRUE)  
attach(data)
```

```
mean(TAS[SEXE==0])
```

```
var(TAS[SEXE==0])
```

```
hist(TAS[SEXE==0], col='blue')
```

```
mean(TAS[SEXE==1])
```

```
var(TAS[SEXE==1])
```

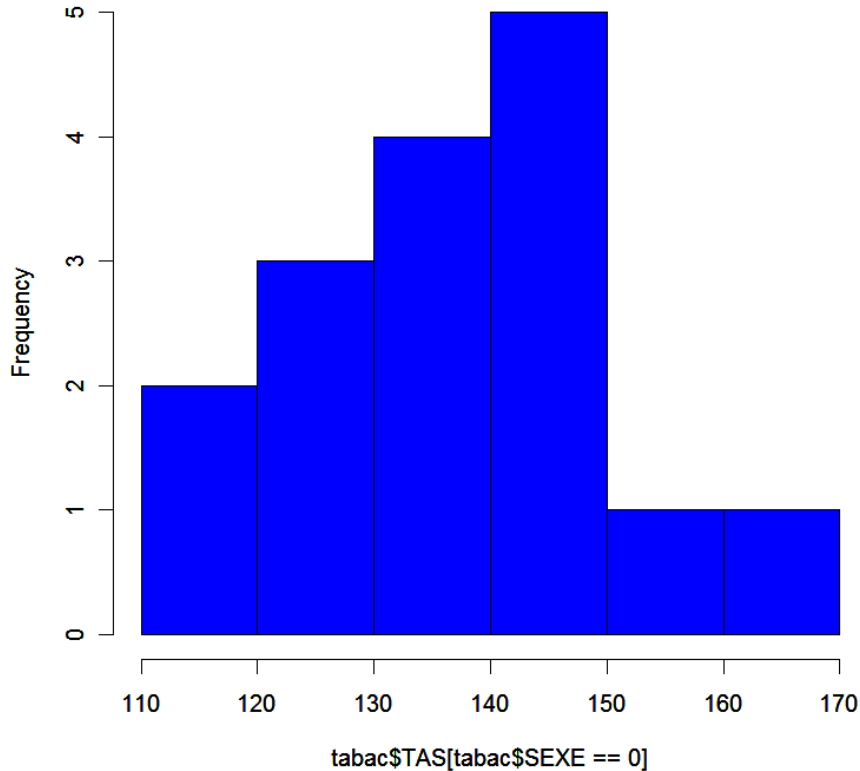
```
hist(TAS[SEXE==1], col='red')
```

```
boxplot(TAS~SEXE)
```

Exercice

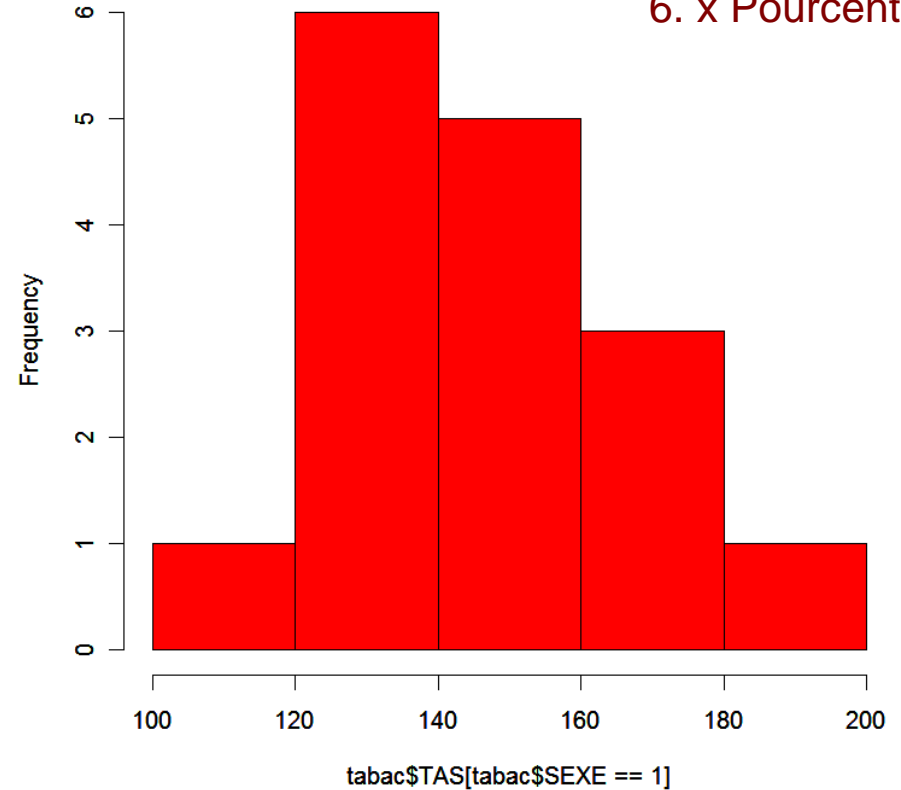
1. Rappels
2. 2 Moyennes
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

Histogram of tabac\$TAS[tabac\$SEXE == 0]



$$m_0 = 136,5 \text{ mmHg}$$
$$s_0^2 = 166,7$$

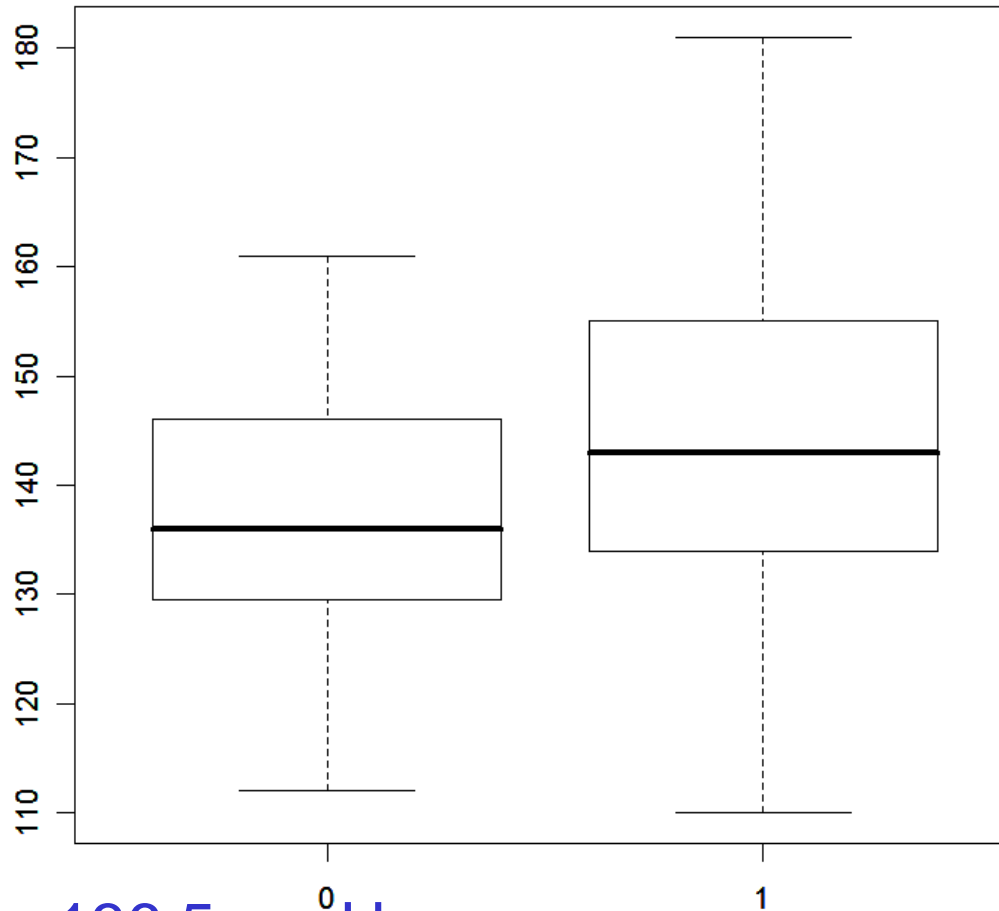
Histogram of tabac\$TAS[tabac\$SEXE == 1]



$$m_1 = 145,1 \text{ mmHg}$$
$$s_1^2 = 316,4$$

Exercice

1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.



$$m_0 = 136,5 \text{ mmHg}$$
$$s_0^2 = 166,7$$

$$m_1 = 145,1 \text{ mmHg}$$
$$s_1^2 = 316,4$$

Exercice

1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

→ 1. Hypothèses

H0: il n'y a pas de différence en moyenne

H1: Il existe une différence de TAS en moyenne entre les hommes et les femmes

→ 2. Prédiction sous H0

Conditions d'application

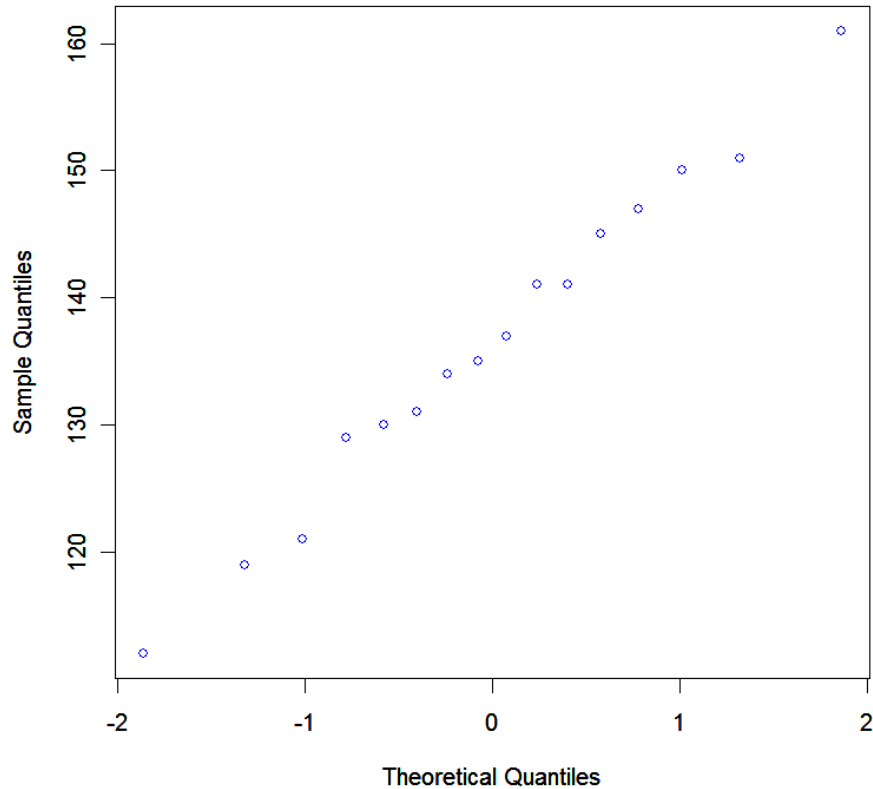
- Normalité ou $n_1 > 30$ et $n_2 > 30$
- Egalité des variance
- Indépendance des individus

Exercice

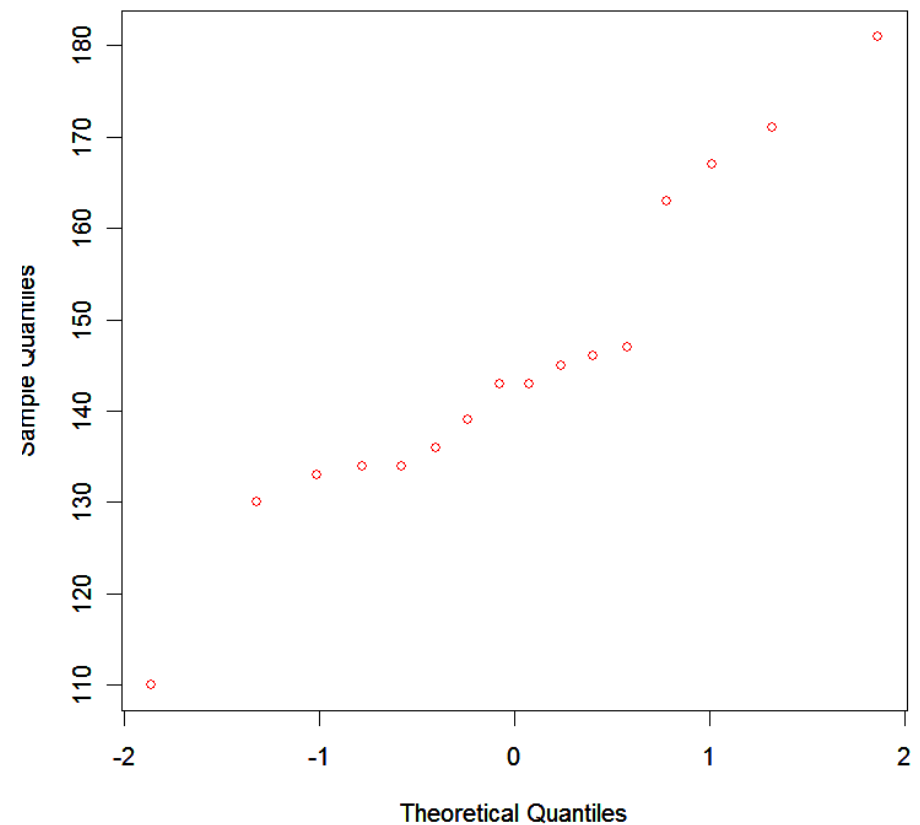
Normalité

1. Rappels
2. 2 Moyennes
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

Normal Q-Q Plot



Normal Q-Q Plot



`qqnorm(TAS [SEXE==0], col="blue")`

`qqnorm(TAS [SEXE==1], col="red")`

Exercice

Variance

var.test(TAS[SEXE==0], TAS[SEXE==1])

1. Rappels
2. 2 Moyennes
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

F test to compare two variances

data: TAS[SEXE == 0] and TAS[SEXE == 1]
F = 0.5268, num df = 15, denom df = 15, p-value = 0.226
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
0.1840566 1.5077135
sample estimates:
ratio of variances
0.5267871

- $p > 0,05$
- Test Non significatif
- Non rejet de H_0 au risque β
- On ne met pas en évidence de différence significative entre les 2 variances

Exercice

1. Hypothèses
2. Prédiction sous H_0

Alors $T \rightarrow T_{\alpha v}$

1. Rappels
- 2. 2 Moyennes**
3. 2 Pourcent.
4. Corrélation
5. x Moyennes
6. x Pourcent.

Exercice

1. Hypothèses
2. Prédiction sous H0
3. Confrontation

t.test(TAS[SEXE==0],TAS[SEXE==1], var.equal=TRUE)

Two Sample t-test

data: TAS[SEXE == 0] and TAS[SEXE == 1]

t = -1.5697, df = 30, p-value = 0.1270

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

95 percent confidence interval:

-19.846469 2.596469

sample estimates:

mean of x mean of y

136.500 145.125

Exercice

1. Hypothèses
2. Prédiction sous H_0
3. Confrontation
4. Interprétation

- $p > 0,05$
- Test Non Significatif
- Non rejet de H_0 au risque β
- On ne met pas en évidence de différence significative pour les TAS, en moyenne, des hommes et de femmes

1. Rappels
2. 2 Moyennes
3. 2 Pourcent
4. Corrélation
5. x Moyennes
6. x Pourcent.

■ Références

- Jean Bouyer: *Méthodes statistiques, Médecine-Biologie*, éditions INSERM
- Coll. (CIMES): *Biostatistiques*, éditions Omnisciences

■ Contact

jean.gaudart@univ-amu.fr

<http://sesstim.univ-amu.fr>

Faculté de Médecine de Marseille