



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*

III. Comparaison de 2 pourcentages observés

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

Exemple: Effet du Tabac sur la survenue de cancer, Echantillon de 32 sujets

15 non fumeurs $P_{nf} = ? \%$

17 fumeurs $P_f = ? \%$

Pourcentage de cancer

table(Tabac, K)

III. Comparaison de 2 pourcentages observés

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

Exemple: Effet du Tabac sur la survenue de cancer, Echantillon de 32 sujets

15 non fumeurs $P_{nf} = ? \%$

17 fumeurs $P_f = ? \%$

Pourcentage de cancer

tab<-table(Tabac,K)

	NK	K
NF	12	3
F	5	12

prop.table(tab,1)

	NK	K
NF	0,8	0,2
F	0,294	0,706

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

→ 1. Hypothèses:

H0: $P_{NF} = P_F$ le pourcentage de cancer est identique chez les fumeurs et les non fumeurs

H1: $P_{NF} \neq P_F$ le pourcentage de cancer est différent chez les fumeurs et les non fumeurs

→ 2. Prédiction:

Sous H0 on doit observer $P = 15/32 = 47\%$ cancers

	NK	K	
NF	12 7,97	3 7,03	15
F	5 9,03	12 7,97	17
	17	15	32

1. Hypothèses

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

→ 2. Prédications

Sous H0 et si les conditions d'application sont respectées

	NK	K	
NF	12 7,97	3 7,03	15
F	5 9,03	12 7,97	17
	17	15	32

$$\chi^2 = \sum \frac{(o_{ij} - c_{ij})^2}{C_{ij}}$$

$$\chi^2 \rightarrow \chi^2_{\nu=1}$$

1. Hypothèses

→ 2. Prédiction

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

Sous H_0 et si les conditions d'application sont respectées

	NK	K	
NF	12 7,97	3 7,03	15
F	5 9,03	12 7,97	17
	17	15	32

Conditions

- $C_{ij} > 5$
- Indépendance des individus

1. Hypothèses

2. Prédiction

→ 3. Confrontation: observation \leftrightarrow théorie sous H_0

chisq.test(Tabac, K, correct=FALSE)

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

1. Hypothèses

2. Prédiction

3. Confrontation: observation \leftrightarrow théorie sous H_0

chisq.test(Tabac, K, correct=FALSE)

Pearson's Chi-squared test

data: Tabac and K

X-squared = 8.1893, df = 1, p-value = 0.004214

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

1. Hypothèses

2. Prédiction

3. Confrontation: observation \leftrightarrow théorie sous H_0

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

chisq.test(Tabac, K, correct=FALSE)

Pearson's Chi-squared test

data: Tabac and K

X-squared = 8.1893, df = 1, p-value = 0.004214

**Test du Chi2 de
Pearson**

1. Hypothèses

2. Prédiction

3. Confrontation: observation \leftrightarrow théorie sous H_0

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

chisq.test(Tabac, K, correct=FALSE)

Pearson's Chi-squared test

data: Tabac and K

X-squared = 8.1893, df = 1, p-value = 0.004214

Données

1. Hypothèses

2. Prédiction

3. Confrontation: observation \leftrightarrow théorie sous H_0

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

chisq.test(Tabac, K, correct=FALSE)

Pearson's Chi-squared test

data: Tabac and K

X-squared = 8.1893, df = 1, p-value = 0.004214

χ^2_0 sous H_0

1. Hypothèses

2. Prédiction

3. Confrontation: observation \leftrightarrow théorie sous H_0

chisq.test(Tabac, K, correct=FALSE)

Pearson's Chi-squared test

data: Tabac and K

X-squared = 8.1893, df = 1, p-value = 0.004214

Petit « p »

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

1. Hypothèses

2. Prédiction

3. Confrontation: observation \leftrightarrow théorie sous H_0

chisq.test(Tabac, K, correct=FALSE)

Pearson's Chi-squared test

data: Tabac and K

X-squared = 8.1893, df = 1, p-value = 0.004214

- $p < 0,05$
- Test significatif
- Rejet de H_0 au risque α
- Il y a une différence entre les 2 pourcentages
- Dans le sens « les fumeurs développent plus souvent de cancer »

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

Conditions d'application:

Où trouver les Cij?

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

```
C<-chisq.test(Tabac, K, correct=FALSE)
```

```
attributes(C)
```

```
$names  
[1] "statistic" "parameter" "p.value" "method" "data.name" "observed"  
[7] "expected" "residuals"  
  
$class  
[1] "htest"
```

```
C$expected
```

	K	
Tabac	0	1
0	7.96875	7.03125
1	9.03125	7.96875

Conditions d'application:

remarque

par défaut

`chisq.test(Tabac, K, correct=TRUE)`

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

Pearson's Chi-squared test with Yates' continuity correction

data: Tabac and K

X-squared = 6.2839, df = 1, p-value = 0.01218

Pour $3 < C_{ij} < 5$



ATTENTION: perte de puissance

1. Hypothèses

2. Prédications

3. Confrontation

→ 4. Interprétation

→ $p < 0,05$

→ Test significatif

→ Rejet de H_0 au risque α

→ Il y a une différence entre les 2 pourcentages

→ Dans le sens « les fumeurs développent plus souvent de cancer »

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

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 le pourcentage de cancer chez les hommes et chez les femmes ?

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 le pourcentage de cancer chez les hommes et chez les femmes ?

tab <- table(SEXE, K)

	NK	K
F	8	8
H	9	7

prop.table(tab, 1)

	NK	K
F	0,5	0,5
H	0,56	0,44

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 entre les hommes et les femmes

H1: il y a une différence dans la survenue du cancer, entre les hommes et les femmes

→ 2. Prédications, conditions d'application

	NK	K
F	8 8,5	8 7,5
H	9 8,5	7 7,5

	NK	K
F	0,5	0,5
H	0,56	0,44

1. Hypothèses

2. Prédiction

3. Confrontation: observation \leftrightarrow théorie sous H_0

chisq.test(SEXE, K, correct=FALSE)

Pearson's Chi-squared test

data: SEXE and K

X-squared = 0.1255, df = 1, p-value = 0.7232

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

1. Hypothèses

2. Prédications

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 entre les 2 pourcentages de cancer

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.

■ 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