



Faculté
de Médecine

Aix-Marseille Université

Tests Non Paramétriques

Plan

1. Paramétriques ou non?
2. Test d'une distribution de probabilité
3. Comparaison de moyennes
4. Comparaison de pourcentages
5. Corrélation

IV. Test de comparaison de pourcentages:

Test exact de Fisher

■ hypothèses:

$H_0: P_1 = P_2$

$H_1: P_1 < P_2$

← Test UNILATERAL

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

■ Principe

	E1	E2	
M+	a	b	m1
M-	c	d	m2
	n1	n2	N

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

- On considère tout les K tableaux possibles
 - ayant les mêmes marges $m1, m2, n1, n2$
 - ayant un écart au moins aussi grand ($\geq |a-b|$), de même sens

- Probabilité d'observer 1 tableau donné, si H0 vraie

$$p_i = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{N!a!b!c!d!}$$

- Probabilité d'observer tableau au moins aussi grand, si H0 vraie:
"p"

$$p = \sum_{i=1}^k p_i$$

■ exemple

Effectif théorique: 2,5

$$p_1 = 0,359$$

obs	E1	E2	
M+	2	3	5
M-	6	5	11
	8	8	16

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

$$|2-3| = 1$$

	E1	E2	
M+	1	4	5
M-	7	4	11
	8	8	16

$$p_2 = 0,128$$

	E1	E2	
M+	0	5	5
M-	8	3	11
	8	8	16

$$p_3 = 0,013$$

$$p = p_1 + p_2 + p_3 = 0,5 > \alpha$$

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

■ Remarques:

- Calculs lourds
- ++ si les effectifs $\uparrow \Rightarrow$ nombre de tableaux \uparrow
- Donne directement " p "
- Test unilatéral \Rightarrow classiquement pour un test bilatéral: $2p$
- possible pour plusieurs pourcentages

■ Condition d'application

- indépendance des individus

Exemple Tabac~Cancer

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

	NK	K
NF	12	3
F	5	12

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

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

1. Hypothèses

2. Prédiction sous H0

la probabilité d'observer un tel tableau sous H0 est grande

3. Confrontation *fisher.test(Tabac,K)*

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélations

1. Hypothèses

2. Prédiction sous H0

la probabilité d'observer un tel tableau sous H0 est grande

3. Confrontation *fisher.test(Tabac,K)*

- 1. Paramétrique?
- 2. Distribution
- 3. Moyennes
- 4. Pourcentages**
- 5. Corrélation

Fisher's Exact Test for Count Data

data: table(Tabac, K)

p-value = 0.00601

alternative hypothesis: true odds ratio is not equal to 1

95 percent confidence interval:

1.506875 71.400728

sample estimates:

odds ratio

8.814559

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

1. Hypothèses

2. Prédiction sous H0

la probabilité d'observer un tel tableau sous H0 est grande

3. Confrontation *fisher.test(Tabac,K)*

Fisher's Exact Test for Count Data

data: table(Tabac, K)

p-value = 0.00601

alternative hypothesis: true odds ratio is not equal to 1

95 percent confidence interval:

1.506875 71.400728

sample estimates:

odds ratio

8.814559

Test

1. Hypothèses

2. Prédiction sous H0

la probabilité d'observer un tel tableau sous H0 est grande

3. Confrontation *fisher.test(Tabac, K)*

- 1. Paramétrique?
- 2. Distribution
- 3. Moyennes
- 4. Pourcentages**
- 5. Corrélation

Fisher's Exact Test for Count Data

data: table(Tabac, K)

p-value = 0.00601

alternative hypothesis: true odds **Données** equal to 1

95 percent confidence interval:

1.506875 71.400728

sample estimates:

odds ratio

8.814559

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

1. Hypothèses

2. Prédiction sous H0

la probabilité d'observer un tel tableau sous H0 est grande

3. Confrontation *fisher.test(Tabac,K)*

Fisher's Exact Test for Count Data

```
data: table(Tabac, K)
```

```
p-value = 0.00601
```

```
alternative hypothesis: true odds ratio is not equal to 1
```

```
95 percent confidence interval:
```

```
1.506875 71.400728
```

```
sample estimates:
```

```
odds ratio
```

```
8.814559
```

« petit p »

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

1. Hypothèses

2. Prédiction sous H0

la probabilité d'observer un tel tableau sous H0 est grande

3. Confrontation *fisher.test(Tabac,K)*

Fisher's Exact Test for Count Data

data: table(Tabac, K)

p-value = 0.00601

alternative hypothesis: true odds ratio is not equal to 1

95 percent confidence interval:

1.506875 71.400728

sample estimates:

odds ratio

8.814559

H1

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

1. Hypothèses

2. Prédiction sous H0

la probabilité d'observer un tel tableau sous H0 est grande

3. Confrontation *fisher.test(Tabac,K)*

Fisher's Exact Test for Count Data

```
data: table(Tabac, K)
```

```
p-value = 0.00601
```

```
alternative hypothesis: true odds ratio is not equal to 1
```

```
95 percent confidence interval:
```

```
1.506875 71.400728
```

```
sample estimates:
```

```
odds ratio
```

```
8.814559
```

Mesure d'association

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

1. Hypothèses

2. Prédiction sous H0

la probabilité d'observer un tel tableau sous H0 est grande

3. Confrontation *fisher.test(Tabac,K)*

Fisher's Exact Test for Count Data

```
data: table(Tabac, K)
```

```
p-value = 0.00601
```

```
alternative hypothesis: true odds ratio is not equal to 1
```

```
95 percent confidence interval:
```

```
1.506875 71.400728
```

```
sample estimates:
```

```
odds ratio
```

```
8.814559
```

IC de la mesure d'association

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 un risque de cancer plus élevée que les non fumeurs”**

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

Exercice SEXE~Cancer

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

	NK	K
F	8	8
H	9	7

	NK	K
F	0,50	0,5
H	0,46	0,54

1. Hypothèses:

Exercice SEXE~Cancer

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

	NK	K
F	8	8
H	9	7

	NK	K
F	0,50	0,5
H	0,46	0,54

1. Hypothèses:

H0: $P_{nf} = P_f$ le pourcentage de cancer est identique chez les hommes et les femmes **OR=1**

H1: $P_{nf} \neq P_f$ le pourcentage de cancer est différent chez les hommes et les femmes **OR≠1**

1. Hypothèses

2. Prédications sous H0

la probabilité d'observer un tel tableau sous H0 est grande

3. Confrontation *fisher.test(SEXE,K)*

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

1. Hypothèses

2. Prédiction sous H0

la probabilité d'observer un tel tableau sous H0 est grande

3. Confrontation *fisher.test(SEXE,K)*

Fisher's Exact Test for Count Data

```
data: table(SEXE, K)
```

```
p-value = 1
```

```
alternative hypothesis: true odds ratio is not equal to 1
```

```
95 percent confidence interval:
```

```
0.1558367 3.8490763
```

```
sample estimates:
```

```
odds ratio
```

```
0.7839342
```

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 lien entre le sexe et la survenue de cancer

1. Paramétrique?
2. Distribution
3. Moyennes
- 4. Pourcentages**
5. Corrélation

1. Paramétrique?
2. Distribution
3. Moyennes
4. Pourcentages
5. Corrélation

■ 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