

## Tests statistiques : pour comparer des fréquences

1

### Plan du cours

- Test statistiques pour fréquences: évaluation de la relation entre deux variables qualitatives
  - *Pour échantillons indépendants*
    - Le test de Khi-deux
    - Le test exact de Fisher
  - *Pour échantillons dépendants*
    - Le test de Mc Nemar
- Exemples d'articles scientifiques et médicales – avec explications
- Exemples des exercices pour l'examen

2

## Le test Khi-deux d'indépendance

- **Nom**
  - **Khi-deux, Khi carre, Khi-2,  $\chi^2$**
  - (ou moins bien chi-deux, chi-2, chi carre)
- **Utilité**
  - pour tester l'**association / relation / lien** entre **deux variables qualitatives**, chaque avec deux ou plus des catégories
  - pour **comparer** des **groups** (2 ou >2) **indépendants concernant** une autre variable **qualitative** avec deux ou plus des catégories
- **Exemples:**
  - Relation entre:
    - *Un facteur de risque et une maladie*
      - Consommation d'alcool (oui/non) et cirrhose (oui/non)
      - Infection VIH (oui/non) et cancer du peau (oui/non)
    - *Un traitement comparé a un autre pour traiter une maladie*
      - Nitroglycérine comparé a placebo pour traiter l'angor (Angine de poitrine ) d'effort stable (oui/non)

3

## Le test Khi-deux d'indépendance

- **Conditions d'application:**
  - deux variables qualitatives (avec deux ou plusieurs valeurs (catégories))
  - observations indépendants dans chaque échantillon
  - échantillons indépendants (on peut considérer que les catégories d'une variable indiquent les échantillons qui sont comparées – eg.: avec sans facteur de risque, traitement 1 ou 2, ...)
  - Plus (>) **80%** des fréquences **théoriques** doivent être supérieurs à **5**
  - **Si < 20 des fréquences théoriques sont <5**
  - **ET** toutes les fréquences **théoriques** sont supérieures à **1** (la règle du W.G.Cochran)
    - **Si** au moins 20% des fréquences théoriques sont inférieurs à 5, ou il y a au moins d'une fréquence inférieure à 1
      - *le test ne peut pas s'appliquer (on utilise le test exact Khi-deux avec correction du Yates ou le test exact de Fisher – voir les diapos suivantes).*
      - *Pour comprendre la règle, voir le tableau théorique dans les diapos suivantes*
    - *Voir la diapo avec le tableau théorique*
    - *On souligne que il s'agit des fréquences théoriques (probables, attendues), pas les fréquences observées.*

4

## Le test Khi-deux d'indépendance

- **Formuler l'hypothèse nulle** (exemples):
- On peut créer plusieurs variantes correctes de l'hypothèse nulle:
  - Est-ce que il y a une association entre le facteur de risque (présent/absent) et la maladie (présente/absente)?
    - Hypothèse nulle (formulations équivalentes):
      - Il n'y a pas de **relation\*** statistiquement significative **entre le facteur de risque et la maladie**
      - Il n'y a pas d' **association\*** statistiquement significative **entre le facteur de risque et la maladie**
      - Il n'y a pas de **lien\*** statistiquement significative **entre le facteur de risque et la maladie**
      - Il n'y a pas de **dépendance \*** statistiquement significative **entre le facteur de risque et la maladie**
      - Il n'y a pas de **différence** statistiquement significative **entre ceux avec et ceux sans maladie** en ce qui concerne la **frequence** du **facteur du risque**
      - Il n'y a pas de **différence** statistiquement significative **entre ceux avec et ceux sans exposition au facteur de risque** en ce qui concerne la **frequence** de la **maladie**.

5

## Le test Khi-deux d'indépendance

- **L'hypothèse nulle** (exemples):
- On peut créer plusieurs variantes correctes des hypothèses:
  - Est-ce que le traitement (oui/non) et le guérison (oui/non) sont liées?
    - Hypothèses équivalentes:
      - Il n'y a pas de **relation** statistiquement significative **entre le traitement et le guérison**
      - Il n'y a pas de **association** statistiquement significative **entre le traitement et le guérison**
      - Il n'y a pas de **lien** statistiquement significative **entre le traitement et le guérison**
      - Il n'y a pas de **dépendance** statistiquement significative **entre le traitement et le guérison**
      - Il n'y a pas de **différence** statistiquement significative **entre ceux avec et ceux sans traitement** en ce qui concerne la **frequence** du **guérison**
      - Il n'y a pas de **différence** statistiquement significative **entre ceux avec et ceux sans guérison** en ce qui concerne la **frequence** du **traitement**.

6

## Le test Khi-deux d'indépendance

- **Formuler l'hypothèse alternative** (exemples):
- On peut créer plusieurs variantes correctes des hypothèses:
  - Est-ce que il y a une relation entre le facteur de risque (présent/absent) et la maladie (présente/absente) ?
    - Hypothèses équivalentes:
      - Il y a un **relation** statistiquement significative entre le facteur de risque et la maladie
      - Il y a un **association** statistiquement significative entre le facteur de risque et la maladie
      - Il y a un **lien** statistiquement significative entre le facteur de risque et la maladie
      - Il y a un **dépendance** statistiquement significative entre le facteur de risque et la maladie
      - Il y a un **différence** statistiquement significative entre ceux avec et ceux sans maladie en ce qui concerne la **frequence** du facteur du risque
      - Il y a un **différence** statistiquement significative entre ceux avec et ceux sans exposition au facteur de risque en ce qui concerne la **frequence** de la maladie.

7

## Le test Khi-deux d'indépendance

- **L'hypothèse alternative** (exemples):
- On peut créer plusieurs variantes correctes des hypothèses:
  - Est-ce que il y a une lien entre le traitement (oui/non) et le guérison (oui/non)?
    - Hypothèses équivalentes:
      - Il y a un **relation** statistiquement significative entre le traitement et le guérison
      - Il y a un **association** statistiquement significative entre le traitement et le guérison
      - Il y a un **lien** statistiquement significative entre le traitement et le guérison
      - Il y a un **dépendance** statistiquement significative entre le traitement et le guérison
      - Il y a un **différence** statistiquement significative entre ceux avec et ceux sans traitement en ce qui concerne la **frequence** du guérison
      - Il y a un **différence** statistiquement significative entre ceux avec et ceux sans guérison en ce qui concerne la **frequence** du traitement.

8

## LE TEST KHI-DEUX D'INDÉPENDANCE

### Exemple

- étude dans lequel on cherche l'effet
  - du tabagisme (facteur du risque – FR)
  - dans l'apparition du cancer du lèvre inférieure (maladie - M).

### Données:

- On a observé 400 sujets parmi lesquels:
  - Parmi 160 qui sont malades (M+) 80 ont aussi le FR présent (FR+)
  - Parmi 240 qui ne sont pas malades (M-) 50 ont aussi le FR présent (FR+)

Ex. structure fichier Excel

	M+ (avec cancer)	M- (sans cancer)	Total
FR+ (fumeurs)	80	50	130
FR- (non fumeurs)	80	190	270
Total	160	240	400

i	Malade	Facteur de risque
1	oui	oui
2	non	non
3	oui	oui
4	oui	non
...	...	...

9

## LE TEST KHI-DEUX D'INDÉPENDANCE

### Question de recherche:

- On cherche à établir si le tabagisme est lié à l'apparition du cancer du lèvre inférieure, ou on peut dire que on cherche à établir si la présence du cancer du lèvre inférieure est liée/associée à le tabagisme, ou on peut dire que on cherche à comparer ceux exposés à ceux sans exposition au tabac du point de vue de la fréquence du cancer de lèvre.

### Notation

- Le tableau de contingence avec les données obtenus après la collection des données est nommé tableau de contingence observé, et les fréquences dans le tableau sont nommées fréquences observées.

	M+ (avec cancer)	M- (sans cancer)	Total
FR+ (fumeurs)	80	50	130
FR- (non fumeurs)	80	190	270
Total	160	240	400

10

## LE TEST KHI-DEUX – LES ÉTAPES

### 1. Formuler les hypothèses:

- $H_0$ : Il n'y a pas un **relation / association / lien** statistiquement significative **entre le facteur de risque (FR) et la maladie (M)** ou Il n'y a pas un **différence** statistiquement significative **entre ceux qui fument et ceux qui ne fument pas la pipe** en ce qui concerne la **fréquence** du **cancer de la lèvre inférieure**.
- $H_1$ : Il y a un **relation / association / lien** statistiquement significative **entre le facteur de risque (FR) et la maladie (M)** ou Il y a un **différence** statistiquement significative **entre ceux qui fument et ceux qui ne fument pas la pipe** en ce qui concerne la **fréquence** du **cancer de la lèvre inférieure**.

### 2. La statistique Khi-deux est:

- On calcule le tableau de contingence théorique en supposant que  $H_0$  est vraie
- On calcule l'écart/ la différence entre le tableau de contingence théorique et le tableau de contingence observée.
- (Si l'écart est:
  - petit,
    - on peut l'expliquer par le hasard,
    - on ne peut pas rejeter l'hypothèse nulle
  - large,
    - c'est plutôt un relation/association/lien qui explique l'écart que le hasard
    - on rejette l'hypothèse nulle)

11

## LE TEST KHI-DEUX D'INDÉPENDANCE

### La réalisation du tableau du contingence théorique/attendu:

- La fréquence théorique d'une cellule = (le total sur la ligne de la cellule  $\times$  le total sur la colonne de la cellule) / le grand total du tableau entière

#### Le tableau observée

	M+	M-	Total
FR+	80	50	130
FR-	80	190	270
Total	160	240	400

#### Le tableau théorique/attendu

	M+	M-	Total
FR+	52	78	130
FR-	108	162	270
Total	160	240	400

On recopie le tableau avec les totaux sauf le contenu!

	M+	M-	Total
FR+	$\frac{130 \times 160}{400}$	$\frac{130 \times 240}{400}$	130
FR-	$\frac{270 \times 160}{400}$	$\frac{270 \times 240}{400}$	270
Total	160	240	400

On peut calculer le reste par différence pour gagner du temps

	M+	M-	Total
FR+	52	130-52	130
FR-	160-52	270-(160-52)	270
Total	160	240	400

12

## LE TEST CHI-2 D'INDÉPENDANCE

Le tableau théorique/attendu est un tableau dans lequel il n'y a pas de relation entre FR et M, donc le pourcentage des malades est le même pour ceux avec ou sans FR, et le pourcentage des sujets avec FR est le même pour ceux avec ou sans maladie

Le tableau observée

	M+	M-	Total
FR+	80	50	130
FR-	80	190	270
Total	160	240	400

Le tableau théorique/attendu

	M+	M-	Total
FR+	52	78	130
FR-	108	162	270
Total	160	240	400

Tableau observée: pourcentages sur lignes

	M+	M-	Total
FR+	62%	38%	130
FR-	30%	70%	270
Total	160	240	400

Tableau théorique: pourcentages sur lignes

	M+	M-	Total
FR+	40%	60%	130
FR-	40%	60%	270
Total	160	240	400

Tableau observée: pourcentages sur colonnes

	M+	M-	Total
FR+	50%	21%	130
FR-	50%	79%	270
Total	160	240	400

Tableau théorique: pourcentages sur colonnes

	M+	M-	Total
FR+	33%	33%	130
FR-	68%	68%	270
Total	160	240	400

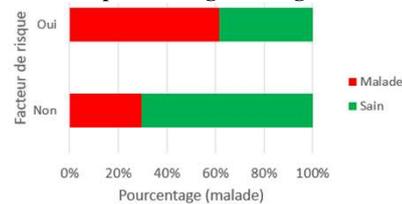
13

## LE TEST CHI-2 D'INDÉPENDANCE

Le tableau théorique/attendu est un tableau dans lequel il n'y a pas de relation entre FR et M, donc le pourcentage des malades est le même pour ceux avec ou sans FR, et le pourcentage des sujets avec FR est le même pour ceux avec ou sans maladie

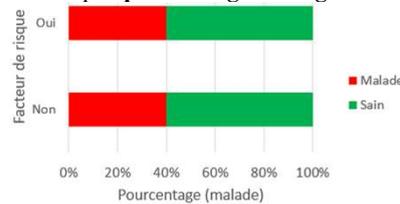
Situation observée

Observée: pourcentages sur lignes

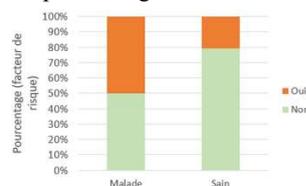


Situation théorique/attendu

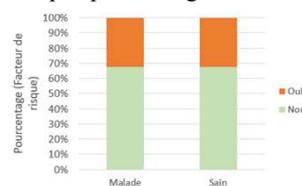
Théorique: pourcentages sur lignes



Observée: pourcentages sur colonnes



Théorique: pourcentages sur colonnes



14

## LE TEST KHI-DEUX D'INDÉPENDANCE

**Le calcul de l'écart / la différence entre le tableau théorique/attendu/prédite et le tableau observée**

**Notation:**

- L – nombres des lignes (des catégories pour la variable FR = 2)
- C – nombres des colonnes (des catégories pour la variable M = 2)
- f – fréquence (nombre/compte), o – observée, t - théorique

**La statistique Khi-deux est**

$$\chi^2 = \sum_{i=1}^{L*C} \frac{(f_i^o - f_i^t)^2}{f_i^t}$$

Le tableau observée

	M+	M-	Total
FR+	80	50	130
FR-	80	190	270
Total	160	240	400

Le tableau théorique/attendu

	M+	M-	Total
FR+	52	78	130
FR-	108	162	270
Total	160	240	400

15

## Les étapes du test Khi-deux

**3. Le niveau de signifiante:  $\alpha = 0,05$**

**4. Les valeurs critiques et la région du rejet:** (voir cours variable aléatoires)

A l'aide de la table de la distribution du Khi-deux: la valeur critique  $X$ , avec  $(L-1)*(C-1)$  degrés de liberté =  $(2-1)*(2-1) = 1 * 1 = 1$  d.d.l.

$X_{(L-1)(C-1); \alpha} = X_{1; 0,05} = 3,84 \Rightarrow RR = [3,84, + \infty)$  – région du rejet

**5. Calculer la valeur de la statistique du test:**

Le tableau observée

	M+	M-	Total
FR+	80	50	130
FR-	80	190	270
Total	160	240	400

Le tableau théorique

	M+	M-	Total
FR+	52	78	130
FR-	108	162	270
Total	160	240	400

$$\chi^2 = \sum_{i=1}^{L*C} \frac{(f_i^o - f_i^t)^2}{f_i^t}$$

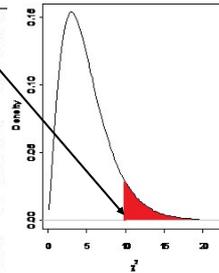
$$X^2 = \frac{(80 - 52)^2}{52} + \frac{(50 - 78)^2}{78} + \frac{(80 - 108)^2}{108} + \frac{(190 - 162)^2}{162} = 37,22$$

16

## Table du distribution Khi-deux

- Ce tableau donne la valeur telle que pour une probabilité alpha  $X^2 > X^2_{v,\alpha}$ , ou  $v = d.d.l$

v	La probabilité $\alpha$									
	0,25	0,2	0,15	0,1	0,05	0,025	0,01	0,005	0,0025	0,001
1	1,3233	1,6424	2,0722	2,7055	3,8415	5,0239	6,6349	7,8794	9,1404	10,8274
2	2,7726	3,2189	3,7942	4,6052	5,9915	7,3778	9,2104	10,5965	11,9827	13,8150
3	4,1083	4,6416	5,3170	6,2514	7,8147	9,3484	11,3449	12,8381	14,3202	16,2660
4	5,3853	5,9886	6,7449	7,7794	9,4877	11,1433	13,2767	14,8602	16,4238	18,4662
5	6,6257	7,2893	8,1152	9,2363	11,0705	12,8325	15,0863	16,7496	18,3854	20,5147
6	7,8408	8,5581	9,4461	10,6446	12,5916	14,4494	16,8119	18,5475	20,2491	22,4575
7	9,0371	9,8032	10,7479	12,0170	14,0671	16,0128	18,4753	20,2777	22,0402	24,3213
8	10,2189	11,0301	12,0271	13,3616	15,5073	17,5345	20,0902	21,9549	23,7742	26,1239
9	11,3887	12,2421	13,2880	14,6837	16,9190	19,0228	21,6860	23,5893	25,4625	27,8767
10	12,5489	13,4420	14,5339	15,9872	18,3070	20,4832	23,2093	25,1881	27,1119	29,5879
11	13,7007	14,6314	15,7671	17,2750	19,6752	21,9200	24,7250	26,7569	28,7291	31,2635
12	14,8454	15,8120	16,9893	18,5493	21,0261	23,3367	26,2170	28,2997	30,3182	32,9092
13	15,9839	16,9848	18,2020	19,8119	22,3620	24,7356	27,6882	29,8193	31,8830	34,5274
14	17,1169	18,1508	19,4062	21,0641	23,6848	26,1189	29,1412	31,3194	33,4262	36,1239
15	18,2451	19,3107	20,6030	22,3071	24,9958	27,4884	30,5780	32,8015	34,9494	37,6978
16	19,3689	20,4651	21,7931	23,5418	26,2962	28,8453	31,9999	34,2671	36,4555	39,2518
17	20,4887	21,6146	22,9770	24,7690	27,5871	30,1910	33,4087	35,7184	37,9462	40,7911
18	21,6049	22,7595	24,1555	25,9894	28,8693	31,5264	34,8052	37,1564	39,4220	42,3119
19	22,7178	23,9004	25,3289	27,2036	30,1435	32,8523	36,1908	38,5821	40,8847	43,8194
20	23,8277	25,0375	26,4976	28,4120	31,4104	34,1696	37,5663	39,9969	42,3358	45,3142



17

## LE TEST KHI-DEUX D'INDÉPENDANCE

6. la décision statistique en fonction de la région du rejet :

- Si  $X$  appartient à  $RR$  nous pouvons rejeter  $H_0$  et accepter  $H_1$
- Si  $X$  n'appartient pas à  $RR$  nous ne pouvons pas rejeter  $H_0$

Parce que:

$X_0 = 37,22$  appartient à  $RR = [3,84, +\infty)$ ,

=> on rejette  $H_0$  avec un risque d'erreur de premier type  $\leq 0,05$ .

La conclusion: nous **acceptons  $H_1$**  avec une erreur de 5% (alpha):

- Il y a un **dépendance** statistiquement significative entre **fumer la pipe** et **le cancer de la lèvre inférieure**
- Il y a un **relation/association/lien** statistiquement significative entre **fumer la pipe** et **le cancer de la lèvre inférieure**
- Il y a un **différence** statistiquement significative entre **ceux avec et ceux sans cancer de la lèvre inférieure** en ce qui concerne la **fréquence** de **fumer la pipe**
- Il y a un **différence** statistiquement significative entre **ceux qui fument et ceux qui ne fument pas la pipe** en ce qui concerne la **fréquence** du **cancer de la lèvre inférieure**.

18

## LE TEST CHI-DEUX D'INDÉPENDANCE

6'. la **décision statistique** en fonction de la **p-value** :

- Si **p-value  $\leq \alpha$  (=0,05) => on rejete  $H_0$  et on accepte  $H_1$** 
  - Il y a une différence/ Il y a une relation – statistiquement significative
- Si **p-value  $> \alpha$  (=0,05) => on ne peut pas rejeter  $H_0$** 
  - On ne peut pas dire qu' Il y a une différence/ On ne peut pas dire qu' Il y a une relation – statistiquement significative

Dans l'exemple: **p < 0,001**

**p < 0,05** nous pouvons **rejeter  $H_0$**  avec un risque d'erreur de premier type  $\leq 0,05$ .

La conclusion: nous **acceptons  $H_1$**  avec une erreur de 5% (alpha):

- Il y a un **dépendance** statistiquement significative entre **fumer la pipe** et **le cancer de la lèvre inférieure**
- Il y a un **relation/association/lien** statistiquement significative entre **fumer la pipe** et **le cancer de la lèvre inférieure**
- Il y a un **différence** statistiquement significative entre ceux avec et ceux sans **cancer de la lèvre inférieure** en ce qui concerne la **fréquence** de **fumer la pipe**
- Il y a un **différence** statistiquement significative entre ceux qui **fument** et ceux qui ne **fument pas la pipe** en ce qui concerne la **fréquence** du **cancer de la lèvre inférieure**.

19

## Indicateurs médicaux a partir du tableau de contingence

Pour interpréter la force de la relation entre un facteur de risque et un maladie on a des indicateurs médicales:

- **RIE - le risque individuel pour les exposées**
  - $RIE = a / (a + b)$  : le % des malades parmi ceux exposées au FR
- **RIN – le risque individuel pour les non exposées**
  - $RIN = c / (c + d)$  : le % des malades parmi ceux non exposées au FR
- **RR – le risque relatif**
  - $RR = RIE / RIN$  : combien du fois le risque d'être malade est plus grand chez les exposées compare a ceux non exposées au FR
- **RA – le risque attribuable** (la différence de risque)
  - $RA = RIE - RIN$  – combien du plus (%) de risque d'être malade a un sujet chez les exposées compare a ceux non exposées au FR

- M – maladie, FR – facteur de risque
- M+ - malade, M- - sans maladie
- F+ - avec facteur de risque
- F- - sans facteur de risque

	M+	M-	Total
FR+	a	b	a+b
FR-	c	d	c+d
Total	a+c	b+d	N=a+b+c+d

20

## Le test Exact Fisher

- **Utilité**
  - pour tester l'**association / relation / lien** entre **deux variables qualitatives**, chaque avec deux ou plus des catégories
  - pour **comparer** des **groups** (2 ou >2) **indépendants concernant** une autre variable **qualitative** avec deux ou plus des catégories
- **Conditions d'application:**
  - deux variables qualitatives (avec deux ou plusieurs valeurs (catégories))
  - observations indépendants dans chaque échantillon
  - échantillons indépendants (on peut considérer que les catégories d'une variable indiquent les échantillons qui sont comparées – eg.: avec sans facteur de risque, traitement 1 ou 2, ...)
  - Il est utilisé quant on ne peut pas utiliser le test Khi-deux.
    - Si Moins de (<) 80% des fréquences **théoriques** sont supérieurs à 5,
    - **Si au moins 20% des fréquences théoriques sont inférieures à 5,**
    - **Si  $\geq 20$  des fréquences théoriques sont  $\leq 5$**

21

## Le test Exact Fisher

- **Hypothèse nulle et alternative – comme pour le test Khi2 !!!**  
**Voir avant!!!**
- **Technique** – est un test qui demande beaucoup des calculs. On ne va pas le calculer nous mêmes. Pour le résultat on va appeler aux ordinateurs.
- Parfois est difficile de calculer le test quand il y a une grand nombre des catégories pour les deux variables
- Le test exact Fisher nous donne presque toujours la bonne valeur du p
- On doit savoir quand on l'applique et comment on interprète le résultat

22

## LE TEST EXACT FISHER

### 6'. la décision statistique en fonction de la p-value :

- Si **p-value  $\leq \alpha$  (=0,05) => on rejete  $H_0$  et on accepte  $H_1$** 
  - Il y a une différence/ Il y a une relation – statistiquement significative
- Si **p-value  $> \alpha$  (=0,05) => on ne peut pas rejeter  $H_0$** 
  - On ne peut pas dire qu' Il y a une différence/ On ne peut pas dire qu' Il y a une relation – statistiquement significative

23

## Test statistique pour comparer caractéristiques qualitatives dichotomiques dans 2 échantillons **appariés**

Test utilisée: test de Mc Nemar

- **Objectif:**

enquête de la signifiante de la différence d'une caractéristique qualitative dichotomique étudiée sur deux échantillons dépendants/ appariés

- **Conditions d'application:**

- observations indépendants dans chaque échantillon
- deux variables qualitative dichotomiques!
- échantillons dépendants/appariés (voir cours introductif tests statistiques)
  - la situation des **mesures répétées** sur les mêmes sujets.
    - ex. on mesure la température d'un patient avant et après la prise d'un médicament qui baisse la température: ibuprofène
  - on compare des **jumeaux identiques**
  - on compare des **échantillons appariés** (pour chaque sujet dans un group on trouve un sujet dans l'autre groupe avec caractéristiques similaires)
  - on compare plusieurs observations faites sur certaines parties du corps du même individu, qui conceptuellement ne sont pas différentes – ex: **partie gauche** avec la **partie droite** (ex. un test dermatologique avec une traitement pour diminuer la perte de cheveux). La comparaison ne doit pas être seulement entre la partie gauche et droite. On peut comparer la partie antérieure et postérieure du scalp.
  - on compare **des méthodes** pour **mesurer** une caractéristique (ou deux techniques diagnostiques [ex tension artérielle]) **sur les mêmes sujets**

- **Hypothèse nulle et alternative – comme pour le test Khi2 !!! Voir avant!!!**

24

24

## Test de Mc Nemar

### Données:

Dans un **étude cas témoin**, dans lequel 63 femmes avec cancer d'endomètre ont été **appariées** avec 63 femmes sans cancer d'endomètre pour être plus similaire possible. Le facteur de risque est la prise d'œstrogènes

Ex: Le **question** est: la prise d'œstrogènes est liée a le cancer d'endomètre?

FR – facteur de risque

a – nb. Cas et témoins qui ont le FR+,

b – nb. Cas qui ont le FR+ et témoins qui ont le FR -

	Témoin + FR + Avec œstrogènes	Témoin + FR – Sans œstrogènes	Total
Cas + FR+ Avec œstrogènes	a	b	a+b
Cas + FR- Sans œstrogènes	c	d	c+d
Total	a+c	b+d	N=a+b+c+d

Ex. structure fichier Excel

i	Facteur de risque	
	Cas	Témoin
1	oui	oui
2	non	non
3	oui	oui
4	oui	non
...	...	...

	Témoin + FR + Avec œstrogènes	Témoin + FR – Sans œstrogènes	Total
Cas + FR+ Avec œstrogènes	27	29	56
Cas + FR- Sans œstrogènes	3	4	7
Total	30	33	63

25

25

## Étapes du test

### 1. Les hypothèses:

- $H_0$ : la prise d'œstrogènes n'est pas liée statistiquement significative a le cancer d'endomètre
- $H_1$ : la prise d'œstrogènes est liée statistiquement significative a le cancer d'endomètre

### 2. La statistique du test McNemar (avec la correction du continuité) est:

$$\chi^2_{1ddl} = \frac{(|b-c| - 0,5)^2}{b+c}$$

		Témoin		Total
		FR +	FR -	
Cas	FR+	a	b	a+b
	FR-	c	d	c+d
Total		a+c	b+d	N=a+b+c+d

- suivre la loi de Khi-deux avec 1 degré de liberté si:  $b + c > 25$

26

## TEST MC NEMAR – LES ÉTAPES

**3. Le niveau de signifiacnce  $\alpha = 0,05$**

**4. Les valeurs critiques et la région critique:**(voir cours variable aléatoires)

On trouve dans le tableau du distribution du Khi-deux la valeur critique X, avec 1 degrés de liberté

$$X_{1;0,05} = 3,84 \Rightarrow RR = [3,84, + \infty)$$

**5. Calculer la valeur de la statistique du test:**

$$\chi^2_{\text{ddl}} = \frac{((29-3|-0,5))^2}{29+3} = 20,32$$

		Témoïn		
		FR +	FR -	Total
Cas	FR+	27	29	56
	FR-	3	4	7
Total		30	33	63

27

## TEST MC NEMAR – LES ÉTAPES

• **6. la décision statistique en fonction de la région du rejet :**

Si **X** appartienne a **RR** nous pouvons **rejeter  $H_0$**  et **accepter  $H_1$**

Si **X** n'appartienne pas a **RR** nous **ne pouvons pas rejeter  $H_0$**

$X = 20,32 > 3,84$  nous pouvons rejeter  $H_0$

avec un risque d'erreur de premier type  $\leq 0,05$ .

$X_0 = 20,32$  appartienne a la région du rejet  $RR = [3,84, + \infty)$ ,

$\Rightarrow$  on rejette  $H_0$ .

La conclusion: nous acceptons  $H_1$ :

Il y a un **relation/association/lien** statistiquement significative **entre le la prise d'œstrogènes et le cancer d'endomètre** avec un risque d'erreur de premier type  $\leq 0,05$

28

## TEST MC NEMAR – LES ÉTAPES

### 6'. la décision statistique en fonction de la p-value :

- Si **p-value  $\leq \alpha$  (=0,05) => on rejete  $H_0$  et on accepte  $H_1$** 
  - Il y a une différence/ Il y a une relation – statistiquement significative
- Si **p-value  $> \alpha$  (=0,05) => on ne peut pas rejeter  $H_0$** 
  - On ne peut pas dire qu' Il y a une différence/ On ne peut pas dire qu' Il y a une relation – statistiquement significative

Dans l'exemple:  $p < 0,001$

$p < 0,05$  nous pouvons rejeter  $H_0$   
avec un risque d'erreur de premier type  $\leq 0,05$ .  
=> on rejette  $H_0$ .

La conclusion: nous acceptons  $H_1$ :

Il y a un **relation/association/lien** statistiquement significative **entre le la prise d'oestrogènes et le cancer d'endomètre** avec un risque d'erreur de premier type  $\leq 0,05$

29

## Indicateurs médicaux a partir du tableau de contingence du test McNemar

Pour interpréter la force de la relation entre un facteur de risque et un maladie on a des indicateurs médicales:

- OR (en anglais: odds ratio)
  - rapport des chances, rapport des cotes
$$OR = b/c = 29/3 = 9,66$$
  - Combien de fois il y a plus des chances d'avoir la maladie
    - chez les exposées compare a les non exposées

		Témoïn		Total
		FR +	FR -	
Cas	FR+	27	29	56
	FR-	3	4	7
Total		30	33	63

30

## Exemples d'articles scientifiques avec tests statistiques

## Chi square test

[Nurs Res Pract.](#) 2012;2012:109251. doi: 10.1155/2012/109251. <http://www.hindawi.com/journals/nrp/2012/109251/>

### Gender Differences in Self-Reported Symptoms of Depression among Patients with Acute Coronary Syndrome.

[Frazier L<sup>1</sup>](#), [Yu E](#), [Sanner J](#), [Liu F](#), [Udtha M](#), [Cron S](#), [Coulter S](#), [Bogaev RC](#).

#### Abstract

This study examined the prevalence of self-reported depressive symptoms and the self reported somatic depressive symptoms as measured by the Beck Depression Inventory-II (BDI-II) among patients hospitalized for acute coronary syndrome (ACS), and explored the impact of gender on both. A convenience sample of 789 adults (248 women and 541 men) was recruited for the study during hospital admission for ACS and participants were screened for self-reported depressive symptoms. BDI-II scores of  $\geq 14$  indicate a moderate level of depressive symptoms and this cut-off score was used to categorize patients into depressed and non-depressed groups. **Pearson chi-square tests for independence (categorical variables)** and t tests for independent samples (continuous variables) were used for gender comparisons. Results showed that depressive symptoms during ACS episodes were different between women and men. Women reported greater overall depressive symptoms (BDI-II mean = 11.89, S.D. = 9.68) than men (BDI-II mean = 9.00, S.D. = 7.93) ( $P < 0.000$ ). Significantly more women (7.66%) were identified positive for somatic depressive symptoms (sleep and appetite disturbances and fatigue) than men (2.22%) ( $P = 0.0003$ ). Findings support that there are gender differences in depressive symptoms experienced by patients hospitalized for ACS. Somatic symptoms of depression may

BDI-II 21 questions symptom present $\geq 1$	All (%)	Females (%)	Males (%)	P value
Q1:Sadness	206 (26.14)	91 (36.69)	115 (21.30)	<0.0001*
Q2:Pessimism	245 (31.05)	86 (34.68)	159 (29.39)	0.1362
Q3:Past failure	174 (22.08)	65 (26.21)	109 (20.19)	0.0583
Q4:Loss pleasure	330 (41.83)	103 (41.53)	227 (41.96)	0.9101

31

## Interprétation du tableau dans l' étude

- On observe que: ils comparent les caractéristique de dépression (présente/absente) pour lequel le score est plus grand que 1 (dans la première cellule - BDI-II 21 questions symptôme présent  $\geq 1$ ), avec le test chi deux (dans le résumé) – en fonction du sexe (les deux colonnes – females /males); pour chaque variable qualitative, on a le nombre des sujets et le pourcentage, résultat du test chi deux (P value)
- Énoncez les hypothèses nulle et alternative, pour le test qui vérifie s'il y a une différence statistiquement significative entre les femmes et les hommes en ce qui concerne la présence du pessimisme (voir dans le tableau les lignes pour Family Type)
- H0 (hypothèse nulle): il n'y a pas une différence statistiquement significative entre les femmes et les hommes en ce qui concerne la présence du pessimisme.
- H1 (hypothèse alternative - test bilatéral): il y a une différence statistiquement significative entre les femmes et les hommes en ce qui concerne la présence du pessimisme.
- Écrivez le nom du test utilisé pour la comparaison:

Le test Khi-deux

32

32

### Interprétation du tableau dans l'étude

Ecrivez la valeur du P du test Khi-deux (voir la colonne P)

- $p=0.1362$

Interpréter du point de vue statistique le résultat du test statistique , et argumentez votre réponse

On peut pas dire qu'il y a une différence statistiquement significative entre les femmes et les hommes en ce qui concerne la présence du pessimisme.

- parce que  $p=0.136$  est plus grand que le niveau de signification de 0.05 (on ne peut pas rejeter l'hypothèse nulle)
- Interpréter du point de vue clinique les résultats:
- **Ecrivez les pourcentages du pessimisme pour chaque groupe**  
Les femmes: 34,68%, et les homes: 29,39%
- **Ecrivez quel groupe a le pourcentage le plus grande**
- Les femmes (voir  $34,68\% > 29,39\%$ )

3  
3

33

### Exemples d'articles scientifiques avec tests statistiques

#### Mc Nemar test

PLoS One. 2012;7(7):e39496. doi: 10.1371/journal.pone.0039496. Epub 2012 Jul 11.

**Vaccination behaviour influences self-report of influenza vaccination status: a cross-sectional study among health care workers.**

Llupia A<sup>1</sup>, García-Basteiro AL, Mena G, Ríos J, Puig J, Bayas JM, Trilla A.

**BACKGROUND:** Published influenza vaccination coverage in health care workers (HCW) are calculated using two sources: self-report and vaccination records. The objective of this study was to determine whether self-report is a good proxy for recorded vaccination in HCW, as the degree of the relationship is not known, and whether vaccine behaviour influences self-reporting.

**METHODS:** A cross-sectional study was conducted using a self-administered survey during September 2010. Considering the vaccination record as the gold standard of vaccination, the properties of self-report as a proxy of the record (sensitivity, specificity, positive predictive value, negative predictive value) were calculated. Concordance between the vaccination campaigns studied (2007-2010) was made using the Kappa index, and **discordance was analyzed using McNemar's test.**

**RESULTS:** 248 HCW responded. The 95% confidence intervals of coverage according to the vaccination record and to self-report overlapped, except for 2007, and the Kappa index showed a substantial concordance, except for 2007. McNemar's test suggested that differences between discordant cases were not due to chance and it was found that the proportion of unvaccinated discordant cases was higher than that of vaccinated discordant cases.

**CONCLUSIONS:** In our study population, self-reported influenza vaccination coverage was a good proxy of the vaccination record. However, vaccination behaviour influences the self-reporting compared to the vaccination record. The sources of coverage should be carefully evaluated and made.

Campaign	Record		Coverage		McNemar P
	Yes	No	True (95% CI)	P	
2007	Declare Yes	28 43	0.19 (0.14-0.24)		$p < 0.0001$
	No	9 118			
2008	Declare Yes	52 18	0.33 (0.27-0.40)		$p = 1$
	No	19 122			
2009s	Declare Yes	71 25	0.32 (0.25-0.37)		$p = 0.002$
	No	7 145			
2009A	Declare Yes	28 13	0.13 (0.09-0.17)		$p = 0.021$
	No	3 204			

34

### Interprétation du tableau dans l' étude

- On observe que: ils comparent les réponses au questionnaire des cadres médicaux s'ils ont été vaccinée (voir sur les lignes – declare yes/no) avec les informations du dossiers de vaccination (voir sur les colonnes – Record yes/no), avec le test Mc Nemar (dans le tableau et le resume); on a le tableau de contingence et le résultat du test Mc Nemar (P)
- Énoncez les hypothèses nulle et alternative, pour le test qui vérifie s'il y a une différence statistiquement significative entre les réponses des cadres médicaux s'ils ont été vaccinée et ceux qui est écrite dans les dossiers de vaccination dans 2007
- H0 (hypothèse nulle): il n'y a pas s'il y a une différence statistiquement significative entre les réponses des cadres médicaux s'ils ont été vaccinée et ceux qui est écrite dans les dossiers de vaccination dans 2007
- H1 (hypothèse alternative - test bilatéral): il y a une différence statistiquement significative entre les réponses des cadres médicaux s'ils ont été vaccinée et ceux qui est écrite dans les dossiers de vaccination dans 2007
- Écrivez le nom du test utilisé pour la comparaison:

Le test Mc Nemar

3  
5

35

### Interprétation du tableau dans l' étude

**Ecrivez la valeur du P du test Khi deux** (voir la colonne P)

- $P < 0.00001$
- **Interpréter du point de vue statistique le résultat du test statistique , et argumentez votre réponse**
- Il y a une différence statistiquement significative entre les réponses des cadres médicaux s'ils ont été vaccinée et ceux qui est écrite dans les dossiers de vaccination dans 2007
- parce que  $p = 0.00001$  est plus grand que le niveau de signification de 0.05 (on rejeté l'hypothèse nulle et on accepte l'hypothèse alternative)

3  
6

36

### Exemples d'articles scientifiques avec tests statistiques

#### Chi square and Fisher exact tests

[J Appl Oral Sci. 2011 May-Jun;19\(3\):212-7. Jaber MA. http://www.scielo.br/so.201100030006&script=sci\\_arttext](#)

**Dental caries experience, oral health status and treatment needs of dental patients with autism.**

**OBJECTIVES:** Autism is a lifelong neurodevelopmental disorder. The aims of this study were to investigate whether children with autism have higher caries prevalence, higher periodontal problems, or more treatment needs than children of a control group of non-autistic patients, and to provide baseline data to enable comparison and future planning of dental services to autistic children.

**MATERIAL AND METHODS:** 61 patients with autism aged 6-16 years (45 males and 16 females) attending Dubai and Sharjah Autism Centers were selected for the study. The control group consisted of 61 non-autistic patients chosen from relatives or friends of autistic patients in an attempt to have matched age, sex and socioeconomic status. Each patient received a complete oral and periodontal examination, assessment of caries prevalence, and caries severity. Other conditions assessed were dental plaque, gingivitis, restorations and treatment needs. **Chi-square and Fisher's exact test** of significance were used to compare groups.

**RESULTS:** The autism group had a male-to-female ratio of 2.8:1. Compared to controls, children with autism had significantly higher decayed, missing or filled teeth than unaffected patients and significantly needed more restorative dental treatment. The restorative index (RI) and Met Need Index (MNI) for the autistic children were 0.02 and 0.3, respectively. The majority of the autistic children either having poor 59.0% (36/61) or fair 37.8% (23/61) oral hygiene compared with healthy control subjects. Likewise, 97.0% (59/61) of the autistic children had gingivitis.

**CONCLUSIONS:** Children with autism exhibit treatment than non-autistic healthy control group importance for children and young people w

**Table 4- Oral hygiene and gingival status of the 61 autistic patients and 61 healthy control subjects**

Treatment needs	Autistic patients		Control		Total	
	No	%	No	%	No	%
Good oral hygiene	2	3.3*	36	59.0	38	31.1
Fair oral hygiene	23	38.0*	16	26.2	39	32.0
Poor oral hygiene	36	59.0*	9	14.8	45	37.0
Gingivitis	59	97.0	25	41.0	84	100
Generalized	46	78.0*	5	20.0	51	61.0
Localized	13	22.0*	20	80.0	33	39.0

**p<0.05 as compared to the control group**

37

## Récapitulatif des tests utilisés

Tests statistiques pour deux échantillons indépendants, variables qualitatives dichotomiques						
Type variable	Nb sujets	Nature des données	Statistique comparée	Test utilise	Paramètre du test	Région du rejet
Qualitative dichotomique	$n_1, p_1 > 10$ , $n_2, p_2 > 10$ , $n_1(1-p_1) > 10$ , $n_2(1-p_2) > 10$		fréquences	Test Z * Equivalent au test Chi carrée ci dessous (Chi carrée est recommandé):	$Z = \frac{P_1 - P_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} - \frac{1}{n_2}\right)}}$ $p = \frac{P_1 n_1 + P_2 n_2}{n_1 + n_2}$	$(-\infty, -Z_{\alpha/2}] \cup [Z_{\alpha/2}, +\infty)$
	<20% cellules du tableau théorique sont <5, et toutes les cellules sont $\geq 1$		fréquences	Test Khi carrée * $v=(L-1)*(C-1) = 1$ d.d.l.	$\chi^2 = \sum \frac{(f_i^o - f_i^t)^2}{f_i^t}$	$[\chi_{v,\alpha}^2, +\infty)$
	<20% cellules du tableau théorique sont <5, et toutes les cellules sont $\geq 1$		fréquences	Test Khi carrée avec correction Yates $v=(L-1)*(C-1)$ d.d.l.	$\chi^2 = \sum \frac{(f_i^o - f_i^t - 0.5)^2}{f_i^t}$	$[\chi_{v,\alpha}^2, +\infty)$
	>20% cellules du tableau théorique sont <5		fréquences	Test exact Fisher	- (test non paramétrique)	-
Tests statistiques pour deux échantillons dépendants/appariées, variables qualitatives dichotomiques						
Qualitative dichotomique	$b+c > 25$		fréquences	Test Mc Nemar $v=1$ d.d.l.	$\chi_{1,d.f.}^2 = \frac{(b-c-0.5)^2}{b+c}$	$[\chi_{v,\alpha}^2, +\infty)$

\* On préfère pour ce cours (et l'examen) le test Chi carrée/deux pour comparer les fréquences, au lieu du test Z pour comparer les fréquences

$p_1, p_2$  – fréquences;  $n_1, n_2$  – nombre des sujets; L et C – nombres des lignes et des colonnes dans le tableau de contingence,  
 $f_i^o$  – fréquence observée,  $f_i^t$  – fréquence théorique; d.d.l. – degrés de liberté;

38

## Récapitulatif des tests utilisés

### Tests statistiques pour plus des deux échantillons (groups) indépendants

Type variable	Nb sujets	Nature des données	Statistique comparée	Test utilisé	Paramètre du test	Région du rejet
Qualitative		<20% cellules du tableau théorique sont <5, et toutes les cellules sont >= 1	Fréquence	Test Khi carrée $v=(L-1)*(C-1)$ d.d.l.	$\chi^2 = \sum \frac{(f_o - f_i)^2}{f_i}$	$[\chi^2_{v, \alpha}]$
		<20% cellules du tableau théorique sont <5, et toutes les cellules sont >= 1	Fréquence	Test Khi carrée avec correction Yates $v=(L-1)*(C-1)$ d.d.l.	$\chi^2 = \sum \frac{(f_o - f_i - 0.5)^2}{f_i}$	$[\chi^2_{v, \alpha}]$
		>20% cellules du tableau théorique sont <5	fréquence	Test exact Fisher	-	-
Quantitative		Normale distribuées, Variance des échantillons égaux	moyenne	Test ANOVA $v1 = \text{nb. groups} - 1$ $v2 = \text{nb. obs.} - \text{nb groups}$	$F = \frac{MCG}{MCE}$	$[F_{v1, v2, \alpha}]$
		Normale distribuées, Variance des échantillons inégaux	moyenne	ANOVA de Welch ou Brown Forsyth		
		Non normale distribuées	médiane	test Kruskal Wallis	-	-

L et C – nombres des lignes et des colonnes dans le tableau de contingence,  $f_o$  – fréquence observée,  $f_i$  – fréquence théorique; d.d.l. – degrés de liberté;

$$MCEG = [n1 * (m1 - mt)^2 + n2 * (m2 - mt)^2 + n3 * (m3 - mt)^2 + \dots] / (p - 1)$$

$$MCDG = [(n1 - 1) * DS1^2 + (n2 - 1) * DS2^2 + (n3 - 1) * DS3^2 + \dots] / (n - p)$$

n=nombre total d'observations, n1, n2, n3 – le nombre d'observations par group, p – nombre des groups

mt – la moyenne des toutes les observations, m1, m2, m3... les moyennes par groups, DS<sub>1</sub>, DS<sub>2</sub>, DS<sub>3</sub>, ... déviations standard d'échantillonnage des groups

39

## Conclusions

- le test Khi-deux, pour comparer 2 échantillons indépendantes, pour vérifier la relation entre deux variables qualitatives.
- Si les fréquences attendues sont petites, le test Khi-deux ne donne pas des bons résultats, donc on doit utiliser le test exact du Fisher, qui offre des meilleurs valeurs du p.
- Le test Mc Nemar pour comparer caractéristiques qualitatives dichotomiques dans 2 échantillons appariées

40

## Exemples des questions pour l'examen

1) \* Une étude a été menée pour vérifier s'il y a un lien statistiquement significatif entre la pratique du cyclisme (oui/non) et les douleurs de genoux (avec/sans douleur). Le tableau de contingence théorique a dans 1 des cellules sur 4 des valeurs plus petites que 5. Le meilleur test statistique pour évaluer le lien est:

- a) le test Khi deux
- b) le test Fisher
- c) le test t pour des échantillons indépendants avec variances égales
- d) le test Bonferonni
- e) le test Bartlett

**Réponse:** b

### Explications

Ici il s'agit de choisir un test statistique à utiliser dans une certaine situation.

On observe que l'énoncé indique un test statistique pour la relation entre deux variables qualitatives [la pratique du cyclisme (oui/non) et les douleurs de genoux (avec/sans douleur) – les deux variables ont comme des valeurs, des catégories, sans nombres => qualitatives. En plus ils ont seulement 2 catégories, donc ils sont dichotomiques].

On doit se poser la question **s'il y a une comparaison** dans la situation.

Parce que on a au moins une variable qualitative, on peut considérer, que l'une des variables qualitatives peut être considérée comme variable de regroupement – elle identifie plusieurs groupes qui sont comparés. Par exemple, on peut considérer la pratique du cyclisme une variable de regroupement – variable pour identifier les groupes à comparer (ce n'est pas obligatoire de choisir la pratique de cyclisme comme variable de regroupement, On peut prendre les douleurs de genoux comme variable de regroupement. Ça dépend de notre préférence, ou on fonction de la modalité de réalisation de l'étude. Mais du point de vue de la choix du test, ça n'importe pas). Comme ça on a deux groupes – ceux qui pratiquent le cyclisme et ceux qui ne pratiquent pas le cyclisme. Donc il y a une comparaison. Si on a dans l'énoncé des groupes, on doit se poser la question: est-ce que les groupes sont dépendantes (appariées) ou indépendantes?

41

## Exemples des questions pour l'examen

### Explications, continuation

On cherche donc dans l'énoncé les 6 situations types des observations dépendantes: mesures répétées sur les mêmes sujets (ex. on mesure la température d'un patient avant et après la prise d'un médicament); comparaison des jumeaux identiques; comparaison des échantillons appariés (pour chaque sujet dans un groupe on trouve un sujet dans l'autre groupe avec caractéristiques similaires); comparaison des plusieurs observations faites sur certaines parties du corps du même individu, qui conceptuellement ne sont pas différentes – ex: partie gauche avec la partie droite (ex. un test dermatologique avec un traitement pour diminuer la perte de cheveux); comparaison des deux (ou plusieurs) méthodes pour mesurer une caractéristique (ou deux techniques diagnostiques [ex tension artérielle]) sur les mêmes sujets. On ne trouve pas aucune de ces situations, donc on considère que les échantillons sont indépendants.

Pour la choix d'un test statistique, s'il s'agit d'une comparaison, on se pose la question, **qu'est qu'il comparent (quelle est le type de la variable/ caractéristique d'intérêt comparée)?** On observe qu'il compare les douleurs de genoux. On a vu que il est qualitative dichotomique. D'une manière implicite, quand on compare des caractéristiques qualitatives, notre but est de comparer leur fréquence d'apparition.

Pour la comparaison des groupes indépendants visant des variables qualitatives, on peut utiliser soit le test Khi-deux, soit le test exact Fisher.

Pour faire la choix entre les deux, on doit savoir quelles sont les fréquences dans le tableau théorique (attendue). La pratique commune est de essayer d'utiliser le test Khi deux, si les conditions d'application du test Khi deux sont satisfaites. Si les conditions d'application ne sont pas satisfaites, on va utiliser le test exact Fisher. En réalité, on peut utiliser toujours le test exact Fisher, et d'ignorer le test Khi-deux. Mais, pour notre cours, et examen, on va faire comme la pratique habituelle, dans la recherche scientifique (premier choix Khi-deux s'il est applicable, sinon Fisher).

42

## Exemples des questions pour l'examen

### Explications, continuation

Si, dans le tableau de contingence théorique, il y a plus de 20% des cellules avec des valeurs inférieure a 5, le test Khi deux n'est pas applicable.

Aussi, Si, dans le tableau de contingence théorique, il y au moins 1 cellule avec des valeurs inférieure égale a 0 ou a 1, le test Khi deux n'est pas applicable.

Dans notre cas on a deux variables qualitatives dichotomiques. Si on fait une tableau de contingence pour la relation entre ces deux variables, on va obtenir 4 cellules.

Dans l'énoncé de la question on peut observer qu'il y a 1 cellule avec des valeurs inférieure a 5. Une cellule sur 4 représente 25% des cellules. 25% est plus grand que 20%. Donc, on a plus de 25% des cellules théoriques avec des valeurs inférieure a 5. Ça indique que le test Khi deux n'est pas applicable. Au final la chois sera le test exact Fisher

43

## Exemples des questions pour l'examen

**2) \* Une étude a été menée pour vérifier s'il y a une relation statistiquement significative entre la consommation régulière du the café (oui/non) et l'apparition d'ostéoporose (présente/absente). Le tableau de contingence attendue a dans toutes les cellules des valeurs plus grandes que 5. Le test statistique qui est le plus utilise pour évaluer ce type des relations est :**

- a) le test Khi deux
- b) le test Fisher
- c) le test t pour des échantillons dépendants
- d) le test Kruskall Wallis
- e) le test F pour les variances

**Réponse:** a

### Explications

Voir exercice 1. La différence est que toutes les valeurs attendues sont plus grandes que 5. Ici la chois sera le test Khi-deux

44

## Exemples des questions pour l'examen

**3) Une étude a été réalisée pour voir s'il y a une relation entre la consommation des fibres et le cancer de colon. La valeur critique du test est 3.84. On a obtenu le tableau suivant des valeurs :**

- |          | cancer + | cancer - |
|----------|----------|----------|
| fibres + | 20       | 980      |
| fibres - | 10       | 990      |
- Lesquelles des suivantes affirmations sont correctes ?
  - a) le paramètre chi deux du test est 3.4
  - b) le paramètre chi deux du test est 3.81
  - c) le paramètre chi deux du test est 0.85
  - d) les valeurs dans les cellules du tableau théorique sont : 15, 985, 15, 985
  - e) les valeurs dans les cellules du tableau théorique sont : 20, 980, 10, 990
  - **Réponse:** a, d

### Explications

**Le calcul de la valeur théorique d'une cellule est: le total qui correspond sur la ligne \* le total qui correspond sur la colonne / grand total.**

**Pour la première cellule  $(20 + 980) * (20 + 10) / (20 + 980 + 10 + 990) = 1000 * 30 / 2000 = 15$**

**Pour la deuxième cellule  $(20 + 980) * (980 + 990) / (20 + 980 + 10 + 990) = 1000 * 1970 / 2000 = 985$**

**Pour la troisième cellule  $(10 + 990) * (20 + 10) / (20 + 980 + 10 + 990) = 1000 * 30 / 2000 = 15$**

**Pour la quatrième cellule  $(10 + 990) * (980 + 990) / (20 + 980 + 10 + 990) = 1000 * 1970 / 2000 = 985$**

45

## Exemples des questions pour l'examen

### 3) Explications, continuation

**Une autre possibilité, est de calculer seulement la première cellule théorique avec la formule présentée, et puis les autres peuvent être calculées par différence avec les totaux**

**Le paramètre du test =  $(20 - 15)^2 / 15 + (980 - 985)^2 / 985 + (10 - 15)^2 / 15 + (990 - 985)^2 / 985 = 3,4$**

46

## Exemples des questions pour l'examen

**4) \* Une étude a été menée pour vérifier s'il y a une relation statistiquement significative entre la consommation régulière du thé café (oui/non) et l'apparition d'ostéoporose (présente/absente). L'étude a été réalisée sur deux échantillons une avec ostéoporose et l'autre sans ostéoporose. Pour chaque sujet avec ostéoporose les chercheurs ont identifié un sujet sans ostéoporose avec le même âge, sexe, et niveau d'effort. Le tableau de contingence attendue a la somme des cellules b + c des valeurs plus grandes que 25. Le meilleur test à utiliser est :**

- a) le test Khi deux
- b) le test exact Fisher
- c) le test t pour des échantillons dépendants
- d) le test Mc Nemar
- e) le test F pour les variances

**Réponse:** d

### Explications

Voir exercice 1. Ici il y a deux échantillons, deux groupes qui sont comparés – ceux avec et sans ostéoporose (l'ostéoporose est la variable de regroupement). L'ostéoporose est dichotomique, aussi que la consommation de café (oui/non, présente/absente). En plus les deux groupes sont appariés (dépendantes), parce que les chercheurs ont identifié un sujet sans ostéoporose avec le même âge, sexe, et niveau d'effort.

Donc il s'agit de comparer deux groupes dépendantes en ce qui concerne une variable dichotomique. Le test à utiliser est le test McNemar. Une condition d'application du test Mc Nemar est que la somme des cellules b + c doit être > 25. Mail l'énoncé nous rassure de cet point de vue.

47

## Exemples des questions pour l'examen

**5) Une étude a été menée pour vérifier s'il y a un lien statistiquement significatif entre la pratique du cyclisme (oui/non) et les douleurs de genoux (avec/sans douleur).**

- a) l'hypothèse nulle est : Il n'y a pas une différence statistiquement significative entre ceux qui pratique le cyclisme et ceux qui ne pratique pas le cyclisme du point de vue de la fréquence des douleurs de genoux
- b) l'hypothèse nulle est : Il n'y a pas une différence statistiquement significative entre ceux avec des douleurs de genoux et ceux sans douleur du point de vue de la fréquence de la pratique du cyclisme
- c) l'hypothèse nulle est : Il n'y a pas une relation statistiquement significative entre la douleur de genoux et la pratique du cyclisme
- d) l'hypothèse nulle est : Il n'y a pas un lien statistiquement significative entre la douleur de genoux et la pratique du cyclisme
- e) l'hypothèse alternative est : Il n'y a pas une relation statistiquement significative entre la douleur de genoux et la pratique du cyclisme

**Réponse:** a, b, c, d

**6) Une étude a été menée pour vérifier s'il y a un lien statistiquement significatif entre la pratique du cyclisme (oui/non) et les douleurs de genoux (avec/sans douleur).**

- a) l'hypothèse nulle est : Il n'y a pas une relation statistiquement significative entre la douleur de genoux et la pratique du cyclisme
- b) l'hypothèse nulle est : Il n'y a pas une dépendance statistiquement significative entre la douleur de genoux et la pratique du cyclisme
- c) l'hypothèse nulle est : Il n'y a pas une association statistiquement significative entre la douleur de genoux et la pratique du cyclisme
- d) l'hypothèse nulle est : Il n'y a pas une différence statistiquement significative entre la douleur de genoux et la pratique du cyclisme
- e) l'hypothèse alternative est : Il y a une relation significative entre la douleur de genoux et la pratique du cyclisme

**Réponse:** a, b, c

**Explications – voila des différents variantes d'exprimer la même chose. Faire attention a les erreurs!**

48

## Exemples des questions pour l'examen

9) Une étude a été réalisée pour voir s'il y a une relation entre la consommation des fibres et le cancer de colon. La valeur critique du test est 3,84. On a obtenu le tableau suivant des valeurs :

	cancer +	cancer -
fibres +	20	980
fibres -	10	990

- Lesquelles des suivantes affirmations sont correctes ?
- a) la région du rejet du test est  $[3,84; +\infty)$
- b) la région du rejet du test est  $[-\infty; -3,84] \cup [3,84; +\infty)$
- c) la région du rejet du test est  $[-\infty; -1,96] \cup [1,96; +\infty)$
- d) les valeurs dans les cellules du tableau attendu sont : 20, 980, 10, 990
- e) les valeurs dans les cellules du tableau observée sont : 20, 980, 10, 990

• **Réponse: a, e**

**Explications:** Il s'agit de comparer deux échantillons (groupes) – consommation des fibres (oui/non), indépendants (aucune information que les échantillons sont dépendants), en ce qui concerne une variable qualitative (cancer de colon – oui/non) => soit le test Khi-deux soit Fisher exact. Si on regarde les fréquences théoriques, toutes sont  $> 5$ , donc on peut utiliser le test Khi-deux.

La région du rejet pour le test Khi deux est [valeur critique,  $+\infty$ )

10) Une étude a été menée pour vérifier s'il y a une relation statistiquement significative entre la consommation régulière du café (oui/non) et l'apparition d'ostéoporose (présente/absente). L'étude a été réalisée sur des jumeaux (avec et sans ostéoporose). Le tableau de contingence attendue a dans toutes les cellules des valeurs plus grandes que 5. :

- a) Le meilleur test à utiliser est le test Chi deux
- b) la région du rejet du test est  $[3,84; +\infty)$
- c) la région du rejet du test est  $[-\infty; -3,84] \cup [3,84; +\infty)$
- d) Le meilleur test à utiliser est le test Mc Nemar
- e) la région du rejet du test est  $[-\infty; -1,96] \cup [1,96; +\infty)$

• **Réponse: b, d**

**Explications:** Un test pour comparer deux groupes – ostéoporose présente/absente – échantillons dépendants, parce que l'étude a été réalisée sur des jumeaux. La comparaison se fait concernant une variable dichotomique => le test McNemar. La région du rejet pour le test McNemar est similaire à Khi deux est [valeur critique,  $+\infty$ )

49

## Exemples des questions pour l'examen

11) Une étude a été réalisée pour voir s'il y a une relation entre la consommation des fibres et le cancer de colon. La valeur critique du test est 3,84. On a obtenu la statistique du test  $X=2$ . Toutes les fréquences théoriques sont  $>5$ .

- Lesquelles des suivantes affirmations sont correctes ?
- a) on rejette l'hypothèse nulle et on accepte l'hypothèse alternative
- b) on rejette l'hypothèse alternative et on accepte l'hypothèse nulle
- c) on ne rejette pas l'hypothèse nulle
- d) il n'y a pas de relation statistiquement significative entre la consommation des fibres et le cancer de colon
- e) on ne peut pas dire qu'il y a de relation statistiquement significative entre la consommation des fibres et le cancer de colon

• **Réponse: c, e**

**Explications:**

Il s'agit de comparer deux échantillons (groupes) – consommation des fibres (oui/non), indépendants (aucune information que les échantillons sont dépendants – voir les 6 situations possibles), en ce qui concerne une variable qualitative (cancer de colon – oui/non) => soit le test Khi-deux soit Fisher exact. Si on regarde les fréquences théoriques, toutes sont  $> 5$ , donc on peut utiliser le test Khi-deux.

La région du rejet pour le test Khi deux est [valeur critique – 3,84,  $+\infty$ )

La statistique du test X est 2, n'appartient pas à la région du rejet, on ne peut pas rejeter l'hypothèse nulle.

La construction du test statistique est autour l'hypothèse nulle, donc l'affirmation b) est fautive.

50

## Exemples des questions pour l'examen

**12) Une étude a été menée pour vérifier s'il y a une relation statistiquement significative entre la consommation régulière du the café (oui/non) et l'apparition d'ostéoporose (présente/absente). L'étude a été réalisée sur des jumeaux (avec et sans ostéoporose). La valeur critique du test est 3,84.  $b+c > 25$ . On a obtenu la statistique du test  $X = 5$ . Lesquelles des suivantes affirmations sont correctes ?**

- a) on rejette l'hypothèse nulle et on accepte l'hypothèse alternative
- b) on rejette l'hypothèse alternative et on accepte l'hypothèse nulle
- c) on ne rejette pas avec l'hypothèse nulle
- d) il y a de association statistiquement significative entre la consommation régulière du the café et l'apparition d'ostéoporose
- e) on ne peut pas dire qu'il y a une relation statistiquement significative entre la consommation des fibres et le cancer de colon

**Réponse:** a, d

**Explications:**

Un test pour comparer deux groupes – ostéoporose présente/absente – échantillons dépendants, parce que l'étude a été réalisée sur des jumeaux.

La comparaison se fait concernant une variable dichotomique => le test McNemar.

La région du rejet pour le test McNemar est similaire à Khi deux est [valeur critique: 3,84, + inf)

La statistique du test  $X$  est 5, appartient pas à la région du rejet, on peut rejeter l'hypothèse nulle, et accepter l'hypothèse alternative.

51

## Exemples des questions pour l'examen

**13) Vous avez trouvé un article scientifique médical qui a étudié les facteurs de risque, les biomarqueurs, l'étiologie et le pronostic de l'AVC ischémique chez les patients cancéreux. L'échantillon d'étude a été partagé en deux groupes: patients ayant l'accident vasculaire cérébral ischémique aigu (AVC) et les patients sans AVC (non AVC). Lesquelles des suivantes affirmations sont correctes ?**

Antécédents médicaux	AVC (n=600)	nonAVC (n=600)	P-valeur
Hypertension artérielle (oui), n (%)	480 (77)	370 (59)	<0,001
Dyslipidémie(oui), n (%)	470 (76)	500 (80)	0,072

- a) les auteurs ont utilisé probablement un test Khi-2 pour des échantillons indépendants
- b) il y a une différence statistiquement significative entre la fréquence de l'AVC chez les patients hypertendus et les patients non hypertendus
- c) on ne peut pas dire qu'il y a une différence statistiquement significative entre la fréquence de l'hypertension artérielle chez les patients avec AVC et les patients non AVC
- d) il y a une association significative entre l'AVC et dyslipidémie
- e) il y a une différence significative entre la fréquence de l'AVC chez les patients ayant dyslipidémie et la fréquence de l'AVC chez les patients sans dyslipidémie

**Réponse:** a, b

52

## Exemples des questions pour l'examen

### 13) Explications:

Il s'agit de comparer deux échantillons (groupes) – avec AVC, et sans AVC, indépendants (aucune information que les échantillons sont dépendants – voir les 6 situations possibles), en ce qui concerne une variable qualitative (hypertension artérielle – oui/non; et puis dyslipidémie – oui/non – le non est implicite) => donc, les auteurs peuvent utiliser soit le test Khi-deux soit le Fisher exact. => réponse a) est correcte. Si les auteurs n'écrivent pas on ne peut pas être sûr. Mais c'est possible, qu'ils ont utilisé l'un des deux. Les fréquences dans le tableau 77 / 600 et 59 / 600 sont très grandes, et ça fait que c'est peu probable que les fréquences théoriques sont < 5. Mais, pour être sûr si les fréquences théoriques sont plus petite que 5 on doit faire le calcul. D'habitude dans autres situations, par exemple, les réactions secondaires, ont des fréquences réduites, et les fréquences théoriques < 5 sont plus probables, et l'utilisation du test exact Fisher est plus fréquente.

Si on regarde la valeur du p, pour la ligne qui contient l'hypertension artérielle, on observe qu'il est < 0,001. Elle est inférieure à 0,05, donc on peut rejeter l'hypothèse nulle, et accepter l'hypothèse alternative. Il y a différentes variantes d'exprimer l'hypothèse alternative, pour le test Khi-deux. On peut considérer comme variable de regroupement, pour faire la comparaison soit le AVC (avec/ sans), soit l'hypertension artérielle (oui/non). Dans la réponse b) ils ont considéré comme variable de regroupement, l'hypertension artérielle. Donc la réponse b) est correcte, et la réponse c) est incorrecte (c'est le contraire).

Si on regarde la valeur du p, pour la ligne qui contient la dyslipidémie, on observe qu'il est 0,072. Elle est supérieure à 0,05, donc on ne peut pas rejeter l'hypothèse nulle. Donc les réponses d) et e) sont incorrectes.

53

## Exemples des questions pour l'examen

### 14) Vous avez trouvé un article scientifique médical qui compare Ivermectin avec Placebo pour prévenir des hospitalisations avec Covid-19. Voici un tableau depuis l'article, et précisez lesquelles des suivantes affirmations sont correctes ?

**Table 2** Outcomes from randomization to end of study visit 95% CI, 95% confidence interval; IQR, interquartile range; MVS: mechanical ventilatory support; SD, standard deviation; # total number of events, \* All adverse events were non-serious

Outcome	Ivermectin (N = 250)	Placebo (N = 251)	Odds Ratio (95% CI)	p value
Primary				
Hospitalization, N (%)	14 (5.60)	21 (8.37)	0.65 (0.32–1.31)	0.227
Secondary				
Time to hospitalization days (in those who were hospitalized), median (IQR)	3.5 (3–5)	3 (2–5)	–	0.59
Invasive MVS, N (%)	4 (1.60)	3 (1.20)	1.34 (0.30–6.07)	0.70
Time to invasive MVS days (in those who required MVS), mean (±SD)	5.25 (1.71)	10 (2)	–	0.019
Negative nasal swab day 3, N (%)	113 (47.08)	120 (49.79)	0.90 (0.63–1.28)	0.55
Negative nasal swab day 12, N (%)	212 (89.08)	221 (92.47)	0.76 (0.45–1.27)	0.29
Dialysis, N (%)	1 (0.40)	1 (0.40)	1.00 (0.06–16.14)	1
All-cause mortality, N (%)	4 (1.60)	3 (1.20)	1.34 (0.30–6.07)	0.70
Safety (adverse events)#, total (%) *	45 (18.00)	53 (21.11)	–	0.6

- Les fréquences absolues du tableau de contingence observée pour le test pour évaluer la différence statistiquement significative entre Ivermectin et Placebo du point de vue de la fréquence du prélèvement nasal négatif jour 3 (Negative nasal swab day 3) sont: 113, 120, 37, 31
- Les fréquences absolues du tableau de contingence attendu pour le test pour évaluer la différence statistiquement significative entre Ivermectin et Placebo du point de vue de la fréquence du prélèvement nasal négatif jour 3 (Negative nasal swab day 3) sont: 116,1; 116,9; 33,9; 34,1
- Les fréquences absolues du tableau de contingence observée pour le test pour évaluer la différence statistiquement significative entre Ivermectin et Placebo du point de vue de la fréquence de la mortalité de toute cause (All-cause mortality) sont: 4, 3, 246, 248

Vallejos L, Zoni A, Barchetti M, Villamandos S, Bobaefilia A, Plano F, Campisi C, Chaperro Campisi E, Medina MF, Achilli F, Guglielmino RA, Ojeda I, Fontana Salazar D, Andino G, Kawerni P, Dellamas S, Aquino AC, Flores V, Martemucci CN, Marinova SM, Sagoni JE, Reynoso P, Sosa NC, Bobledo MC, Quarracena JM, Nunez MM, Babi Diaz N, Mesa E, Aguirre MG. Ivermectin to prevent hospitalizations in patients with COVID-19 (IVERCON-COVID19) a randomized, double-blind, placebo-controlled trial. BMC Infect Dis. 2021 Jul 2;21(1):638. doi: 10.1186/s12879-021-06346-5

54

## Exemples des questions pour l'examen

14)

- d) Les fréquences absolues du tableau de contingence **théorique** pour le test pour évaluer la différence statistiquement significative entre Ivermectin et Placebo du point de vue de la fréquence de la mortalité de toute cause (All-cause mortality) sont: 3,5; 3,5; 246,5; 247,5
- e) 50% des fréquences absolues du tableau de contingence attendu pour le test pour évaluer la différence statistiquement significative entre Ivermectin et Placebo du point de vue de la fréquence de la mortalité de toute cause (All-cause mortality) sont inférieures à 5
- f) Un bon test statistique pour la différence entre Ivermectin et Placebo du point de vue de la fréquence de la mortalité de toute cause (All-cause mortality) est le test Fisher exact
- g) Un bon test statistique pour la différence entre Ivermectin et Placebo du point de vue de la du prélèvement nasal négatif jour 3 est le test Khi deux

Réponse: a, b, c, d, e, f, g

55

## Exemples des questions pour l'examen

14)

**Table 2** Outcomes from randomization to end of study visit 95% CI: 95% confidence interval; IQR: interquartile range; MVS: mechanical ventilatory support; SD: standard deviation; # total number of events, \* All adverse events were non-serious

Outcome	Ivermectin (N = 250)	Placebo (N = 251)	Odds Ratio (95% CI)	p value
Primary				
Hospitalization, N (%)	14 (5.60)	21 (8.37)	0.65 (0.32–1.31)	0.227
Secondary				
Time to hospitalization days (in those who were hospitalized), median (IQR)	3.5 (3–5)	3 (2–5)	–	0.59
Invasive MVS, N (%)	4 (1.60)	3 (1.20)	1.34 (0.30–6.07)	0.70
Time to invasive MVS days (in those who required MVS), mean (±SD)	5.25 (1.71)	10 (2)	–	0.019
Negative nasal swab day 3, N (%)	113 (47.08)	120 (49.79)	0.90 (0.63–1.28)	0.55
Negative nasal swab day 12, N (%)	232 (89.08)	221 (82.47)	0.76 (0.45–1.27)	0.29
Dialysis, N (%)	1 (0.40)	1 (0.40)	1.00 (0.06–16.14)	1
All-cause mortality, N (%)	4 (1.60)	3 (1.20)	1.34 (0.30–6.07)	0.70
Safety (adverse events)#, total (%) *	45 (18.00)	53 (21.11)	–	0.6

	Ivermectin	Placebo	Total
Negative nasal swab day 3 (oui)	113	120	
Negative nasal swab day 3 (non)	250-113=37	251-120=31	
Total	250	251	

### Explications:

a) Pour trouver les fréquences observées, on prend les totaux et on le copie dans le tableau observée, après on prend la fréquence de la caractéristique du tableau Nasal swab day qui nous indique le nombre des sujets qui ont eu un test de prélèvement nasal négatif. Pour trouver la fréquence des ceux qui ont eu un test de prélèvement nasal positive, on fait la différence avec le total de sujets par group. Donc les fréquences observées sont 113, 120, 37, 31.

b) On va suivre les instruction du cours pour calculer le tableau de fréquences théorique (n'oubliez pas théorique = attendue)

56

## Exemples des questions pour l'examen

14)

**Table 2** Outcomes from randomization to end of study visit 95% CI: 95% confidence interval; IQR: interquartile range; MVS: mechanical ventilatory support; SD: standard deviation; # total number of events. \* All adverse events were non-serious

Outcome	Ivermectin (N = 250)	Placebo (N = 251)	Odds Ratio (95% CI)	p value
<b>Primary</b>				
Hospitalization, N (%)	14 (5.60)	21 (8.37)	0.65 (0.32–1.31)	0.227
<b>Secondary</b>				
Time to hospitalization days (in those who were hospitalized), median (IQR)	3.5 (3–4)	3 (2–5)	–	0.59
Invasive MVS, N (%)	4 (1.60)	3 (1.20)	1.34 (0.30–6.07)	0.70
Time to invasive MVS days (in those who required MVS), mean ( $\pm$ SD)	5.25 (1.71)	10 (2)	–	0.019
Negative nasal swab day 3, N (%)	113 (47.08)	120 (49.79)	0.90 (0.63–1.28)	0.55
Negative nasal swab day 12, N (%)	212 (89.08)	221 (92.47)	0.76 (0.45–1.27)	0.29
Dialysis, N (%)	1 (0.40)	1 (0.40)	1.00 (0.06–16.14)	1
All-cause mortality, N (%)	4 (1.60)	3 (1.20)	1.34 (0.30–6.07)	0.70
Safety (adverse events)#, total (%) *	45 (18.00)	53 (21.11)	–	0.6

	Ivermectin	Placebo	Total
All-cause mortality (oui)	4	3	
All-cause mortality (non)	250-4=246	251-3=248	
Total	250	251	

### Explications:

e) Pour trouver les fréquences observées, on prend les totaux et on le copie dans le tableau observée, après on prend la fréquence de la caractéristique du tableau All-cause mortality qui nous indique le nombre des sujets qui ont eu un décès (toute cause). Pour trouver la fréquence des ceux qui ont n'ont pas décédée, on fait la différence avec le total de sujets par group. Donc les fréquences observées sont 4, 3, 246, 248.

b) On va suivre les instruction du cours pour calculer le tableau de fréquences théorique (n'oubliez pas théorique = attendue)

57

## Exemples des questions pour l'examen

14) **Explications:**

e) Deux cellules, parmi 4 cellules (sans les totaux) du tableau de contingence attendu pour le test pour évaluer la différence statistiquement significative entre Ivermectin et Placebo du point de vue de la fréquence de la mortalité de toute cause (All-cause mortality), donc 50%

f) Un bon test statistique pour la différence entre Ivermectin et Placebo du point de vue de la fréquence de la mortalité de toute cause (All-cause mortality) est le test Fisher exact, parce que, il y a plus que 20% des cellules qui ont des valeurs inférieures à 5 (en effet il y a 50% des cellules avec des valeurs inférieure à 5), donc le test Khi deux n'est pas applicable dans cette situation, et on peut utiliser le test exact du Fisher.

g) Un bon test statistique pour la différence entre Ivermectin et Placebo du point de vue de la du prélèvement nasal négatif jour 3 est le test Khi deux, parce que il y a moins que 20% des cellules qui ont des valeurs inférieures à 5 (en effet il y a 0% des cellules avec des valeurs inférieure à 5), donc le test Khi deux est applicable dans cette situation

58

FIN...

Objectifs qui méritent une visite en Roumanie, proche de Cluj – 1 heure

**La Mine de sel de Turda**

<http://salinaturda.eu/?lang=en>



Google street view:

<http://virtualglobetrotting.com/map/salina-turda-salt-mine/view/google/>