

Le test de Chi carré (χ^2)

À quoi ça sert ?

Le test de Chi Carré (prononcer Khi) est un test utiliser en biologie afin de déterminer l'indépendance ou non d'une série de valeur sur une autre.

Dans le cadre du cours nous utiliserons ce test afin de déterminer si deux espèces semble présenter une interdépendance ou non.

Pour faire le test il faut :

Des échantillons aléatoires indépendants.

Des effectifs théoriques (valeurs) qui sont supérieurs ou égaux à 5 individuellement

Un nombre total d'effectifs (valeurs) supérieurs à 30

Ce que nous allons faire c'est comparé les valeurs obtenues avec les valeurs que nous pensions obtenir selon l'équation suivante :

$$\chi_{calculé}^2 = \sum_{j=1}^J \frac{(O_j - E_j)^2}{E_j}$$

Où

O_j = fréquence observé (valeur expérimentale)

E_j = fréquence envisagé (valeurs attendue)

j = Numéro de la catégorie de la valeur

J = Nombre total de catégorie de valeur

Par la suite il est possible d'interpréter la valeur obtenue selon la table suivante :

Tableau d'interprétation de la valeur Chi

Résultat du test Chi entre 2 valeurs	0,02	0,15	0,46	1,1	1,6	2,7	3,8	6,6	7,9
Probabilité de l'hypothèse	90%	70%	50%	30%	20%	10%	5%	1%	0,50%
	Acceptable						Inacceptable		

Considérant que les scientifiques trouvent mathématiquement acceptable une théorie qui obtient un résultat de plus de 5% nous pouvons donc dire si oui ou non il y a une différence significative dans les données.

Exemple 1:

Vous lancez 50 pièces au sol.

De la main gauche :
Résultat 28 piles, 22 faces.

De la main droite :
Résultat 10 piles, 40 faces.

Avez-vous une main avec laquelle vous seriez capable de contrôler les résultats?

Premièrement il faut poser l'hypothèse nul : Que non il ne devrait pas avoir de différence.
Donc 50 : 50

Pour la main gauche
hyp: Non il n'y a pas de différence.
Donc 25:25

$$\chi^2 = \frac{(28-25)^2}{25} + \frac{(22-25)^2}{25}$$

piles faces
 $\frac{9}{25} + \frac{9}{25}$

$\chi^2 = 0,72$ qui est acceptable.

Pour la main droite
hyp: Non il n'y a pas de différence.
Donc 25:25

$$\chi^2 = \frac{(10-25)^2}{25} + \frac{(40-25)^2}{25}$$

$\chi^2 = 18$

Qui est inacceptable!
Donc oui la droite pourrait influencer significativement les résultats.

Exemple 2 :

Deux tigres dans un zoo s'accouplent et ont 4 petits : 2 blancs et 2 normaux.

En sachant que les chances d'obtenir des tigres blancs dû à un gène récessif sont de 25%

Est-ce normal ? (Dû à un gène récessif ou autre chose)

hyp: Non il n'y a pas de différence significative

Donc 25 : 75
pour 4 1 : 3

$$\chi^2 = \frac{(2-1)^2}{1} + \frac{(2-3)^2}{3}$$

$\chi^2 = 1,33$ qui est acceptable

Exercices

1. Au cours d'une étude cas-témoin réalisée dans un hôpital, 317 patientes souffrant d'un cancer de l'endomètre ont été appariées à 317 patientes saines. La prise d'œstrogènes durant les 6 mois précédent le diagnostique a été déterminée. On a observé que 54 femmes saines et 152 femmes qui ont développé le cancer avaient pris des œstrogènes. Faites l'analyse de ces proportions.

634 patients totaux
206 malades
32,44% de malades

hyp: la prise d'œstrogène n'influence pas
Donc proportion égale de malade.
 $317 \cdot 32,44\% = 103$

$$\chi^2 = \frac{(54 - 103)^2}{103} + \frac{(152 - 103)^2}{103}$$

$\chi^2 = 46,62$ l'hypothèse est inacceptable, il y a une différence significative
l'œstrogène semble être un facteur.

2. Au cours d'un essai clinique, 184 personnes ont reçu le médicament traditionnel et 103 personnes ont reçu un nouveau médicament. Avec le médicament traditionnel, on a observé 129 guérisons et avec le nouveau 80. Faites l'analyse des proportions.

287 patients
209 guérisons
72,8% de guérisons

hyp: la prise n'influence pas
Donc proportion égale
 $184 \cdot 72,8 = 134$
 $103 \cdot 72,8 = 75$

$$\chi^2 = \frac{(129 - 134)^2}{134} + \frac{(80 - 75)^2}{75}$$

$\chi^2 = 0,52$ qui est acceptable. Il n'y a donc pas de différence significative entre les 2 résultats.

3. 2 000 personnes sont suivies pendant 20 ans, 800 sont fumeurs et 1200 non fumeurs. Au cours des 20 ans de suivi, on observe 100 cas de cancer: 90 chez les fumeurs, 10 chez les non fumeurs. Effectuez l'analyse statistique.

2000 suivis
100 cancers
→ 5% proportion

hyp: fumer n'a pas d'effet sur le taux de cancers
Donc proportions égales
fumeur 800 · 5% = 40
non-fumeur 1200 · 5% = 60

$$\chi^2 = \frac{(90 - 40)^2}{40} + \frac{(10 - 60)^2}{60}$$

$\chi^2 = 104$ l'hypothèse est inacceptable, fumer semble avoir un effet sur le % cancers.

4. Lors d'une étude sur la pollution bactérienne, la présence ou l'absence de salmonelles a été recensée à partir d'échantillons d'eau prélevés dans 3 bassins européens. Les analyses fournissent les résultats suivants :

	Rhin	Loire	Seine
Présence de salmonelles	8	10	16
Absence de salmonelles	2	11	23

% de pollution 80% 45% 42%

Peut-on affirmer que la Seine est moins polluée que le Rhin et la Loire?

[source: http://www.obs-vlfr.fr/~enseigne/maitp6/solution_test/testexo16.htm]

Loire - Seine
 $\chi^2 = \frac{(42 - 40)^2}{40} = 0,75$ hyp acceptable, pas de différence, donc on ne peut pas affirmer que la Seine est moins polluée.

Rhin - Seine
 $\chi^2 = \frac{(42 - 80)^2}{80} = 18,05$ hyp non acceptable on peut affirmer qu'il y a une différence

5. Un vétérinaire recense 4 cas de brucellose dans un gros élevage extensif de moutons du Larzac comptant 1230 têtes. Ce résultat invalide-t-il de façon significative le modèle épidémiologique affirmant que la maladie ne touche en principe qu'un individu sur 1000?

théor 1 : 1000

Elevage 1,23 = 1230

hyp: il n'y a pas de différence significative

$$\chi^2 = \frac{(4 - 1,23)^2}{1,23}$$

$\chi^2 = 6,238$ l'hypothèse est inacceptable, il semble que le fermier a un problème

6. En général, on enregistre 15 naissances jumellaires sur 1000 naissances. En Suède, une étude a été menée sur 30000 femmes enceintes et 840 femmes ont donné naissances à des jumeaux. Peut-on considérer que la Suède est significativement "hors normes"?

E = 15 : 1000 → 1,5%

real = 840 : 30000 → 2,8%

hyp: il n'y a pas de différence significative

$$\chi^2 = \frac{(2,8\% - 1,5\%)^2}{1,5\%}$$

$\chi^2 = 1,12$ l'hypothèse est confirmée, il n'y a pas de différence significative.