TEST DU χ^2

Le test du χ^2 (khi-carré ou khi-deux) peut être utilisé pour comparer un pourcentage observé à un pourcentage théorique, ou deux ou plusieurs pourcentages observés sur des échantillons indépendants..Comme tous les tests statistiques, il est fondé sur **l'hypothèse nulle**, à savoir l'absence de différence entre pourcentage observé et pourcentage théorique, ou entre les pourcentages observés.

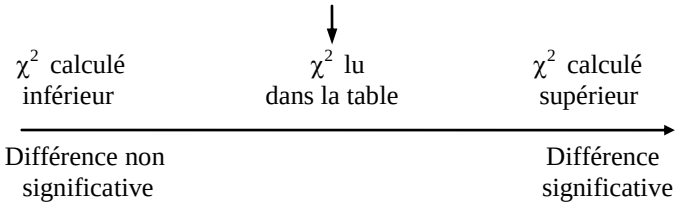
Un indice est calculé, qui quantifie l'écart entre les effectifs observées et les effectifs théoriques (attendus si l'hypothèse nulle est vraie). Sous certaines conditions, cet indice suit une loi de χ^2 vraie. Cette loi décrit les fluctuations d'échantillonnage de cet indice sous l'hypothèse nulle.

La table des lois de χ^2 permet de déterminer la plage des valeurs de cet indice pour des valeurs de probabilité données ou de déterminer la probabilité pour que l'indice χ^2 égale ou dépasse une valeur donnée, lorsque l'hypothèse nulle est vérifiée.

Si l'on prend, comme souvent, la limite de 5 p. cent, on compare la valeur du χ^2 pour 5 p. cent et la valeur de l'indice calculé avec les résultats observés (*cf.* figure χ^2.1).

Figure χ^2.1

Principe de l'interprétation du test du χ^2



Si l'indice calculé est supérieur à la valeur du χ^2 pour 5 p. cent, ceci signifie que l'écart observé a une probabilité inférieure à 5% d'être observée sous l'hypothèse nulle , ;par conséquent, cet écart est jugé peu probable, l'hypothèse nulle est rejetée (par rapport au seuil de risque d'erreur consenti de 5 p. cent), et on conclut que la différence est « statistiquement significative ». Au contraire, si l'indice calculé est inférieur à la valeur du χ^2 pour 5 p. cent, on conclut qu'il n'y a pas de différence significative car l'écart est jugé probable sous l'hypothèse nulle (probabilité supérieure à 5%.d'observer un écart au moins aussi important).

Plus l'indice calculé est élevé et supérieur à la valeur du χ^2 pour 5 p. cent, plus la distance entre des données calculées et les données théoriques est importante et plus la probabilité que cette distance ou différence ne soit due qu'au hasard des fluctuations d'échantillonnage est faible. On peut alors préciser jusqu'à quel faible niveau de probabilité la différence est « significative ».

Nous prendrons successivement des exemples pour :

- Comparer un pourcentage observé à un pourcentage théorique,
- Comparer deux pourcentages observés,
- Comparer plusieurs pourcentages observés.

I - COMPARAISON D'UN POURCENTAGE OBSERVE ET D'UN POURCENTAGE THEORIQUE

Dans une région, le pourcentage connu d'exploitations infectées est de 20. Une étude faite sur un échantillon de 200 exploitations de petite taille tirées au sort a révélé que 24 d'entre elles étaient infectées.

On s'interroge pour savoir si la proportion d'exploitations de petite taille infectées est différente de celle des autres exploitations.

A partir de ces données, on peut établir le tableau χ^2.1

Tableau χ^2.1

Effectif observé et effectif calculé d'élevages infectés et d'élevages indemnes

	Elevages infectés	***Elevages indemnes***	***Total***
Effectif observé O	*24*	*176*	*200*
Effectif calculé C	*40*	*160*	*200*

Ce tableau permet de faire apparaître les effectifs observés (24 exploitations infectées, 176 exploitations indemnes) et les effectifs calculés (C), c'est-à-dire attendus sous l'hypothèse nulle le pourcentage dans les petites exploitations est identique à celui de l'ensemble des exploitations.

L'indice utilisé pour le χ^2 correspond à la somme des quotients : carrés des différences divisés par les effectifs calculés, soit :

$$\chi^2 = \sum \frac{(O-C)^2}{C}$$

Dans le cas présent, le χ^2 vaut :

$$\chi^2 = \frac{(24-40)^2}{40} + \frac{(176-160)^2}{160} = \frac{(16)^2}{40} + \frac{(16)^2}{160} = 6{,}4 + 1{,}6 = 8$$

Ce chiffre est à comparer avec ceux de la table de χ^2 (tableau χ^2.2).

Il faut toutefois, auparavant, préciser le « nombre de degrés de liberté » (d.d.l.) (cf. encadré χ^2.1).

Encadré χ^2.1

Les degrés de liberté

On appelle «*degré de liberté*» le nombre de cases du tableau des effectifs observés (case(s) «total» mise(s) à part) que l'on doit **remplir librement** pour pouvoir reconstituer l'ensemble du tableau à l'aide des marges : il constitue un reflet de la dimension du tableau.

❒ Lorsque l'on examine un échantillon selon **un seul caractère**, le degré de liberté est égal au **nombre d'éventualités** de ce caractère, **moins une**.

1ER EXEMPLE : un caractère (le sexe), avec deux éventualités (mâle ou femelle)

	Mâles	**Femelles**	**Total**
Effectifs			100

Ainsi, dans ce premier exemple, d'un caractère (le sexe), avec deux éventualités (mâle ou femelle), sur un total (quel qu'il soit) de 100 animaux, le degré de liberté est **UN**, c'est-à-dire que l'on peut **fixer librement** l'effectif **d'une seule case**, car lorsque l'on a fixé le nombre de mâles ou de femelles, l'effectif de l'autre case est automatiquement déterminé.

Dans ce cas, le degré de liberté (d.d.l.) est :

d.d.l. = nombre d'éventualités moins une

d.d.l. = 2 - 1 = 1

2EME EXEMPLE : un caractère (le type d'élevage), avec trois éventualités (élevage laitier, allaitant ou mixte)

3	**Elevages laitiers**	**Elevages allaitants**	**Elevages mixtes**	**Total**
Effectifs				100

Dans ce deuxième exemple, il est possible de **remplir librement deux cases** du tableau, mais pas une troisième.

Le degré de liberté est donc de deux, ce qui correspond bien à la règle énoncée ci-dessus :

d.d.l. = nombre d'éventualités moins une
d.d.l. = 3 - 1 = 2

❒ Lorsque l'on examine un échantillon selon **deux caractères**, le nombre de degrés de liberté est égal :

- Au **nombre d'éventualités** du premier caractère, **moins une**,
- **Multiplié** par le **nombre d'éventualités** du deuxième caractère, **moins une**.

On peut illustrer ceci à l'aide d'un troisième exemple.

3EME EXEMPLE : un caractère (le type d'élevage) avec trois éventualités (cf. ci-dessus) et un deuxième caractère (le département) avec deux éventualités (département A et département B).

Eventualités élevage / **Eventualités département**	**Elevages laitiers**	**Elevages allaitants**	**Elevages mixtes**	**Total**
Département A				50
Département B				150
Total	30	70	100	200

Dans un tel tableau, on peut **remplir librement deux cases**, **n'importe lesquelles**. Mais dès que les effectifs de deux cases sont fixés, il n'existe plus aucune liberté car les effectifs de toutes les autres cases sont automatiquement déterminés.

Le nombre de degrés de liberté est donc de deux, ce qui correspond bien à la règle citée ci-dessus :

d.d.l. = (nombre d'éventualités, moins une) (nombre d'éventualités, moins une)
d.d.l. = (3 –- 1) (2 – 1) = 2

Dans le cas présent, pour 2 éventualités (infecté, non infecté), le nombre de d.d.l. est : 2 – 1 = 1

On consulte donc la première ligne de la table de χ^2 (tableau χ^2.2).

On y lit la probabilité pour que le χ^2 égale ou dépasse une valeur donnée. Ainsi, pour un degré de liberté, il y a 50 p. cent de chances pour que l'écart dû au hasard corresponde à une valeur de χ^2 de 0,45, 20 p. cent pour 1,64 et 5 p. cent pour 3,84.

Tableau χ^2.2

La table du χ^2 donne la probabilité *p* du risque α pour que χ^2 égale ou dépasse une valeur donnée, en fonction du nombre de degrés de liberté (d.d.l.)

pα / d.d.l.	0,90	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,0158	0,455	1,074	1,642	2,706	3,841	5,412	6,635	10,827
2	0,211	1,386	2,408	3,219	4,605	5,991	7,824	9,210	13,815
3	0,584	2,366	3,665	4,642	6,251	7,815	9,837	11,345	16,266
4	1,064	3,357	4,878	5,989	7,779	9,488	11,668	13,277	18,467
5	1,610	4,351	6,064	7,289	9,236	11,070	13,388	15,086	20,515
6	2,204	5,348	7,231	8,558	10,645	12,592	15,033	16,812	22,457
7	2,833	6,346	8,383	9,803	12,017	14,067	16,622	18,475	24,322
8	3,490	7,344	9,524	11,030	13,362	15,507	18,168	20,090	26,125
9	4,168	8,343	10,656	12,242	14,684	16,919	19,679	21,666	27,877
10	4,865	9,342	11,781	13,442	15,987	18,307	21,161	23,209	29,588
11	5,578	10,341	12,899	14,631	17,275	19,675	22,618	24,725	31,264
12	6,304	11,340	14,011	15,812	18,549	21,026	24,054	26,217	32,909
13	7,042	12,340	15,119	16,985	19,812	22,362	25,472	27,688	34,528
14	7,790	13,339	16,222	18,151	21,064	23,685	26,873	29,141	36,123
15	8,547	14,339	17,322	19,311	22,307	24,996	28,259	30,578	37,697
16	9,312	15,338	18,418	20,465	23,542	26,296	29,633	32,000	39,252
17	10,085	16,338	19,511	21,615	24,769	27,587	30,995	33,409	40,790
18	10,865	17,338	20,601	22,760	25,989	28,869	32,346	34,805	42,312
19	11,651	18,338	21,689	23,900	27,204	30,144	33,687	36,191	43,820
20	12,443	19,337	22,775	25,038	28,412	31,410	35,020	37,566	45,315
21	13,240	20,337	23,858	26,171	29,615	32,671	36,343	38,932	46,797
22	14,041	21,337	24,939	27,301	30,813	33,924	37,659	40,289	48,268
23	14,848	22,337	26,018	28,429	32,007	35,172	38,968	41,638	49,728
24	15,659	23,337	27,096	29,553	33,196	36,415	40,270	42,980	51,179
25	16,473	24,337	28,172	30,675	34,382	37,652	41,566	44,314	52,620
26	17,292	25,336	29,246	31,795	35,563	38,885	42,856	45,642	54,052
27	18,114	26,336	30,319	32,912	36,741	40,113	44,140	46,963	55,476
28	18,939	27,336	31,391	34,027	37,916	41,337	45,419	48,278	56,893
29	19,768	28,336	32,461	35,139	39,087	42,557	46,693	49,588	58,302
30	20,599	29,336	33,530	36,250	40,256	43,773	47,962	50,892	59,703

*Si, comme d'habitude, la probabilité retenue est de 5 p. cent, l'écart entre les effectifs observés et les effectifs calculés ne doit pas égaler ou dépasser **3,84** sous l'hypothèse nulle (c'est-à-dire qu'il n'y a pas de différence entre la proportion d'exploitations de petite taille infectées et la proportion des exploitations infectées de la région).*

Dans le cas présent, on a trouvé un χ^2 de 8, c'est-à-dire supérieur à 3,84. On peut donc rejeter l'hypothèse nulle, avec un risque d'erreur de 5 p. cent, et admettre que la différence entre le pourcentage d'exploitations de petite taille infectées et le pourcentage d'exploitations infectées dans la région est significative.

CONDITIONS D'APPLICATION DU χ^2

Dans tous les cas, le test du χ^2 ne peut être utilisé que si deux conditions sont remplies. Elles sont rappelées dans l'encadré χ^2.2, commentées et illustrées ci-dessous.

Encadré χ^2.2

Les conditions d'application du χ^2

1. Le test du χ^2 ne peut être utilisé que si les effectifs calculés de toutes les cases sont d'au moins 5.
2. Les « individus » dénombrés doivent être indépendants sur le plan statistique.

Première condition

Ici la condition était respectée puisque les effectifs calculés étaient de 40 et de 160.

Il arrive qu'une ou plusieurs cases d'un tableau de contingence corresponde(nt) à des effectifs théoriques inférieurs à 5.

- Le test de χ^2 demeure applicable lorsque plus de 80 p. cent des cases ont des effectifs théoriques supérieurs ou égaux à 5. Ainsi, pour un tableau de contingence avec **plusieurs degrés de liberté** (par exemple, 3x2, soit deux degrés de liberté), comportant seulement une case dont l'effectif théorique est inférieur à 5 (et donc 5 cases sur 6 avec des effectifs théoriques supérieurs à 5, soit 5/6 = 83 p. cent), il est possible d'utiliser le test de χ^2. Si plus de 20 p. cent des cases comprennent des effectifs théoriques inférieurs à 5, il faut utiliser le test exact de Fisher.
- Pour un tableau de contingence à **un seul degré de liberté** (tableau 2x2), il est possible d'utiliser la **correction de Yates** lorsqu'un seul des effectifs théoriques est inférieur à 5 (mais supérieur à 3). Elle consiste à retrancher 0,5 à chaque différence absolue 'effectif observé-effectif théorique', avant de l'élever au carré :

$$\chi^2 = \sum \frac{(|O - C| - 0{,}5)^2}{C}$$

Lorsque l'effectif théorique est plus faible (≤ 3), il faut utiliser le test exact de Fisher.

Deuxième condition

L'indépendance des « individus » sur le plan statistique signifie qu'il ne doit pas y avoir de lien entre eux, notamment au travers d'une appartenance à des groupes d'individus (par exemple, en médecine vétérinaire, tous les animaux d'un même élevage ont en commun le fait d'être soumis aux mêmes conditions d'élevage, au même environnement ; ils ne sont pas indépendants par rapport aux animaux d'autres élevages).

Exemple :

Soit une étude faite sur l'administration d'antibiotiques lors du tarissement des vaches, avec 20 élevages dans lesquels on applique le tarissement systématique et 20 élevages dans lesquels on ne l'applique pas.

L'utilisation du test du χ^2 sur un tableau de contingence fondé sur l'unité « élevage » (avec, par exemple, un taux de mammites supérieur ou inférieur à 5 p. cent) est possible, car les élevages sont indépendants les uns des autres.

En revanche, elle ne le serait pas pour un tableau de contingence fondé sur l'unité « animal » car les vaches de chaque élevage ne sont pas indépendantes les unes des autres.

N.B. Le test du χ^2 fournit les mêmes résultats que le test de l'écart-réduit. Ces deux tests sont identiques, le χ^2 étant le carré de ε pour un degré de liberté. La différence réside dans le fait que l'on peut utiliser le test du χ^2 lorsqu'il existe plusieurs degrés de liberté (**comparaison de plusieurs pourcentages**, cf. plus loin), alors qu'on ne le peut pas avec le test de l'écart-réduit.

II - COMPARAISON DE DEUX POURCENTAGES OBSERVES

On peut reprendre l'exemple 1 utilisé dans la présente annexe, concernant le nombre d'exploitations infectés trouvées dans deux échantillons issus de deux régions A et B.

Région A : 200 exploitations dont 40 infectées
Région B : 300 exploitations dont 30 infectées

A partir de ces données, on peut bâtir le tableau χ^2.3.

Tableau χ^2.3

Effectifs observés dans deux régions et effectifs calculés correspondants (entre parenthèses)

	Exploitations infectées		Exploitations indemnes		Total
Région A	40	(28)	160	(172)	200
Région B	30	(42)	270	(258)	300
Total	70		430		500

Les effectifs observés apparaissent dans le tableau.

Pour le calcul des effectifs théoriques, on va utiliser l'hypothèse nulle selon laquelle les pourcentages d'exploitations infectées sont identiques dans la région A et la région B. Dans ce cas, on obtient 70 exploitations infectées sur un total de 500 exploitations, soit un pourcentage de 100 x 70/500 = 14.

Dans la région A, le nombre « calculé » d'exploitations infectées attendu sous l'hypothèse nulle est de :

$$200 \times 14 \text{ p. cent} = 28$$

placé entre parenthèses dans la case supérieure gauche.

Dans la région B, par analogie, le nombre « calculé » d'exploitations infectées est de :

$$300 \times 14 \text{ p. cent} = 42$$

placé entre parenthèses dans la case inférieure gauche.

Les effectifs calculés d'exploitations indemnes s'obtiennent par soustraction et sont placés entre parenthèses dans les cases de droite.

D'une façon générale, l'effectif calculé (ou attendu sous H_o) pour une case donnée peut être obtenu en multipliant le total de la ligne correspondante par le total de la colonne correspondante et en divisant par le total général du tableau.

Le calcul du χ^2 est donc le suivant :

$$\chi^2 = \frac{(40 - 28)^2}{28} + \frac{(160 - 172)^2}{172} + \frac{(30 - 42)^2}{42} + \frac{(270 - 258)^2}{258} = 9{,}97$$

Le nombre de d.d.l. pour un tableau de ce genre est le produit :

$$(\text{nombre de colonnes} - 1)\ (\text{nombre de lignes} - 1)$$

$$\text{Soit, ici : } (2 - 1)\ (2 - 1) = 1$$

Dans la table de χ^2, le seuil de signification à 5 p. cent pour un d.d.l. est $\chi^2 = 3{,}84$.

La valeur trouvée : 9,97, dépasse cette limite et permet donc de considérer que l'écart observé dans les proportions d'exploitations infectées des régions A et B est **significatif** et probablement non dû simplement aux fluctuations d'échantillonnage.

La valeur trouvée : 9,97, est d'ailleurs supérieure à celle correspondant à un risque α de 1 p. cent (cf. table χ^2, ligne 1 : 6,635) et inférieure à celle correspondant à un risque α de 1 p. mille (10,827).

La différence entre les deux régions A et B est donc significative avec un degré de signification compris entre 1 p. cent et 1 p. mille. Il aurait été possible de dire que la différence constatée dans les pourcentages d'exploitations infectées des deux régions A et B est significative par rapport au seuil de risque de 1 p. cent si ce seuil avait été choisi plutôt que celui de 5 p. cent.

Enfin, on peut signaler que la condition d'utilisation du test du χ^2 a bien été respectée puisque les effectifs calculés étaient tous supérieurs à 5.

III - COMPARAISON DE PLUSIEURS POURCENTAGES OBSERVES

Dans une zone, on a effectué une enquête descriptive d'une maladie et on a estimé le pourcentage d'exploitations infectées dans trois catégories d'exploitations (strates) :

- Nombre d'animaux de 1 à 20 ;
- Nombre d'animaux de 21 à 50 ;
- Nombre d'animaux supérieur à 50.

Les résultats obtenus figurent dans le tableau χ^2.4.

Tableau χ^2.4

Effectifs observés en fonction de la taille des exploitations et effectifs calculés

	Exploitations				
	infectées		**indemnes**		**Total**
Exploitations de 1 à 20 animaux	10	(20)	90	(80)	100
Exploitations de 21 à 50 animaux	12	(12)	48	(48)	60
Exploitations de plus de 50 animaux	18	(8)	22	(32)	40
Total	40		160		200

Comme dans le cas précédent, le calcul des effectifs théoriques fait appel à l'hypothèse nulle et conduit aux chiffres placés entre parenthèses dans le tableau χ^2.4. Dans ce cas, on constate qu'il est nécessaire de calculer au moins deux effectifs avant d'en déduire les autres par soustraction (2d.d.l.).

Le calcul du χ^2 se présente comme suit :

$$\chi^2 = \frac{(10-20)^2}{20} + \frac{(90-80)^2}{80} + \frac{(12-12)^2}{12} + \frac{(48-48)^2}{48} + \frac{(18-8)^2}{8} + \frac{(22-32)^2}{32} = 21{,}87$$

Le nombre de degrés de liberté est : $(3-1)(2-1) = 2$.

La consultation de la table du χ^2 pour 2 d.d.l. ($2^{ème}$ ligne) montre que la valeur obtenue au seuil de 5 p. cent est de 5,99. Les pourcentages d'exploitations infectées dans les trois catégories d'exploitations classées en fonction de la taille diffèrent donc de façon significative puisque la valeur observée (21,87) est supérieure à la valeur 5,99. On peut donc dire que les pourcentages d'infection des exploitations sont différents selon la taille des exploitations.

Le degré de signification (c'est-à-dire le plus petit risque d'erreur pour lequel la différence est encore significative) est inférieur à 1 p. mille (cf. table de χ^2 : tableau χ^2.2) : χ^2 pour un risque α de 0,001 pour 2 d.d.l. = 13,815. Il aurait été possible de conclure à un écart significatif par rapport à un seuil de 1 p. mille s'il avait été choisi plutôt que celui de 5 p. cent.

Par ailleurs, comme précédemment, on aura vérifié que la condition d'utilisation du test de χ^2 était respectée puisque chaque effectif calculé est supérieur à 5.

IV - DEGRE DE SIGNIFICATION ET FORCE DU REJET DE L'HYPOTHESE NULLE

Il est tentant de prendre en considération le degré de signification comme support indicateur de la faiblesse du risque de se tromper en rejetant l'hypothèse nulle : plus le risque alpha est faible, et plus on serait conforté dans le rejet d'Ho.

Ce serait une erreur.

Pour s'en convaincre, il suffit tout d'abord de considérer qu'en augmentant suffisamment les effectifs… n'importe quel écart peut devenir significatif !

Le degré de signification doit être utilisé uniquement pour savoir de façon **qualitative** si le seuil est dépassé, pour conclure alors à un écart significatif et pour rien d'autre !

D'autant que l'indicateur pertinent pour juger de la « force » de la validité du choix de l'hypothèse H1 (du fait du rejet de Ho), c'est la puissance, qui relève du risque béta, et non du risque alpha (voir Annexe : Les erreurs de jugement).