

MODULE 6 : Deux variables sont-elles liées ?

CAPSULE 2 : LE KHI2 CALCULÉ ET SON INTERPRÉTATION

Lorsque l'on revient sur la façon dont on a calculé notre khi2, on peut relever une observation importante. En effet, le chiffre obtenu a plus de chances d'être élevé à mesure que le tableau augmente... puisqu'il y aura plus de cellules dans le tableau. Cet élément doit être pris en compte pour interpréter le khi2. Il s'agit de ce que l'on nomme le « degré de liberté ».

Pourquoi ce nom de « degré de liberté » (noté dli) ? Parce que plus vous avez des cellules au sein d'un tableau, moins la variation d'une cellule interviendra sur les autres. Regardez, c'est assez simple. Dans le tableau qui suit, il suffit de connaître la valeur d'une seule cellule pour en déduire les autres. Le degré de liberté est donc de 1.

	PLUTÔT ENTRETENUE	PLUTÔT NATURELLE	TOTAL
FORÊT 1 COMBE DE L'IF	DONNÉE 1 = 12	?	30
FORÊT 2 COL DE PORTE	?	?	30
TOTAL	28	32	60

DEGRÉ DE LIBERTÉ = 1

Dans cet autre tableau, on comprend qu'il est nécessaire d'avoir deux informations pour en déduire toutes les autres. Le degré de liberté est donc de 2.

	VOITURE	VÉLO	TOTAL	TOTAL
HOMME	DONNÉE 1	DONNÉE 2	?	9
FEMME	?	?	?	12
TOTAL	7	8	6	21

DEGRÉ DE LIBERTÉ = 2

Il existe une formule simple pour calculer le degré de liberté d'un tableau.

$d_{ll} = (\text{nb de lignes} - 1) \times (\text{nb de colonnes} - 1)$ où le nombre de lignes et de colonnes s'entend sans les lignes ou colonnes de total.

Alors, quel est le degré de liberté de notre tableau 1 ?

TABLEAU 1 – tris croisé sentiment de bonheur/niveau de revenus en effectifs – situation observée

TABLEAU 1 : TRIS CROISÉS SENTIMENT DE BONHEUR/NIVEAU DE REVENUS EN EFFECTIFS - SITUATION OBSERVÉE

		SENTIMENT DE BONHEUR				TOTAL
		TRÈS HEUREUX 😊	PLUTÔT HEUREUX 🙂	PAS TRÈS HEUREUX 😐	PAS HEUREUX DU TOUT 😞	
3 LIGNES {	JUSQU'À 1900 EUROS 📄	44	172	47	8	271
	DE 1901 À 3100 EUROS 📄	63	197	16	0	276
	PLUS DE 3100 EUROS 📄	81	192	6	0	279
	TOTAL	188	561	69	8	826

4 COLONNES

DEGRÉ DE LIBERTÉ = (3-1) X (4-1) = 6

KHI2 CALCULÉ = 69

$DII = (3-1) \times (4-1) = 6$. Notre degré de liberté est de 6 pour un khi2 calculé de 69.

Afin d'être interprété, notre khi2 calculé doit être comparé à un khi2 théorique, dont la valeur varie en fonction du degré de liberté et des marges d'erreurs. Tous ces calculs théoriques sont rassemblés dans une table, la table du khi2. Il serait fastidieux de vous expliquer comment est calculée cette table. Il suffit pour nous de savoir qu'elle existe et que c'est avec elle que nous allons interpréter le khi2. En voici un extrait.

EXTRAIT DE LA TABLE DU KHI2

TABLE DU χ^2

La table donne la probabilité p pour que χ^2 égale ou dépasse une valeur donnée, en fonction du nombre de degrés de liberté dll

dll	0,99	0,975	0,95	0,90	0,10	0,05	0,025	0,01
1	0,0002	0,001	0,004	0,016	2,71	3,84	5,02	6,63
2	0,02	0,05	0,10	0,21	4,61	5,99	7,38	9,21
3	0,11	0,22	0,35	0,58	6,25	7,81	9,35	11,34
4	0,30	0,48	0,71	1,06	7,78	9,49	11,14	13,28
5	0,55	0,83	1,15	1,61	9,24	11,07	12,83	15,09
6	0,87	1,24	1,64	2,20	10,64	12,59	14,45	16,81
7	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48

DEGRÉS DE LIBERTÉ

PROBABILITÉS D'ERREUR

69
DONC
 $p < 0,1\%$

En haut de la table figurent les probabilités d'erreur (notées « p »). 0,10 signifie 10% de probabilité d'erreur.

Sur la gauche figurent les degrés de liberté (dll).

Reportons-nous à la ligne 6, qui correspond au degré de liberté, et voyons où placer notre khi2 calculé. Le chiffre 69 ne figure pas sur cette ligne, qui s'arrête au chiffre de 16,749. Cela signifie que notre khi2 calculé sort de la table par la droite, et a donc une probabilité d'erreur qui est située sous le seuil de 0,01 (soit 1%).

Autrement dit, on peut être certains à plus de 99% que nos deux variables ne sont pas indépendantes.

Que se serait-il passé si nous avions eu un khi2 inférieur ? En sociologie, le seuil de probabilité admis est de 5% (0,05). Au dessus, on ne peut écartier l'hypothèse de l'indépendance des variables. Si notre khi2 calculé avait été inférieur au khi2 théorique de 12,591, il nous aurait été impossible d'écartier l'hypothèse de l'indépendance des variables avec une précision suffisante.

Dit plus simplement : si votre Khi2 se situe à gauche de la colonne 0,05, vous ne pouvez pas interpréter votre tableau sans prendre de risques.

Remarquez que plus le degré de liberté diminue, plus les khi2 théoriques diminuent. Cela revient à dire qu'en réduisant le nombre de lignes et/ou de cellules d'un tableau, il est possible de valider le test du khi2. On nomme cette opération « recomposer un tableau ».

EXTRAIT DE LA TABLE DU KHI2

TABLE DU χ^2

La table donne la probabilité p pour que χ^2 égale ou dépasse une valeur donnée, en fonction du nombre de degrés de liberté dll

dll	0,99	0,975	0,95	0,90	0,10	0,05	0,025	0,01
1	0,0002	0,001	0,004	0,016	2,71	3,84	5,02	6,63
2	0,02	0,05	0,10	0,21	4,61	5,99	7,38	9,21
3	0,11	0,22	0,35	0,58	6,25	7,81	9,35	11,34
4	0,30	0,48	0,71	1,06	7,78	9,49	11,14	13,28
5	0,55	0,83	1,15	1,61	9,24	11,07	12,83	15,09
6	0,87	1,24	1,64	2,20	10,64	12,59	14,45	16,81
7	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48

KHI2