



MIASS 241
Mathématiques (appliquées aux sciences sociales) 4
© El Hadj Touré, 2023

NEUF EXERCICES RÉCAPITULATIFS- Chi-carré
(Solutionnaire)

1) Comparez, en termes de différences et de similitudes, les paires de concepts ci-dessous :

a) Analyse de tableaux croisés et test du chi-carré

L'analyse de tableaux croisés et le test du chi-carré sont deux techniques qui reposent sur les tableaux de distribution de fréquences et dont l'enjeu est d'analyser la relation entre deux variables catégorielles ou de type discret.

Toutefois, l'analyse de tableaux croisés demeure approximative puisqu'elle repose sur la comparaison des différences entre deux catégories de la VI eu égard à la VD. De surcroît, elle porte uniquement sur des données d'échantillon. Par contre, le test du chi-carré est plus précis en ce sens qu'il intègre dans son calcul toute l'information fournie par le tableau croisé. De plus, il permet d'examiner dans quelle mesure la relation établie à l'aide de l'analyse tabulaire bivariée s'avère statistiquement significative au niveau de la population dont provient l'échantillon.

NB : Si l'analyse tabulaire bivariée repose habituellement sur des % (fréquences relatives), le test du chi-carré est toujours basé sur des fréquences (absolues) 😊

b) Test du chi-carré et mesure d'association comme le coefficient de contingence,

Le test du chi-carré et la mesure d'association qu'est le coefficient de contingence sont deux techniques permettant d'analyser la **relation entre deux variables catégorielles**, c.à.d. de type qualitatif ou discret. Elles ont ceci de commun qu'elles mettent en évidence dans quelle mesure une relation s'avère significative.

Toutefois, le test du chi-carré permet d'établir la **signification statistique** d'une relation entre deux variables catégorielles alors que les mesures d'association révèlent plutôt la **signification réelle** d'une relation.

NB : De petites différences peuvent facilement être significatives du point de vue statistique, lorsque le nombre de cas est très grand (n). En conséquence, après avoir établi la signification statistique d'une relation (généralisation), il faut toujours en préciser la signification réelle (intensité) 😊

c) Niveau de confiance et seuil de signification statistique

Le niveau de confiance réfère au degré de certitude considéré dans l'**estimation par intervalle** d'un paramètre de la population à l'aide d'une statistique d'un échantillon (95% par exemple). Le seuil de signification statistique, quant à lui, renvoie à la probabilité théorique de commettre une erreur en rejetant l'hypothèse nulle dans un **test d'hypothèse** (0,05 ou 5% par exemple).

Toutefois, le niveau de confiance et le seuil de signification réfèrent tous au **degré de précision** que l'on désire conférer aux résultats d'échantillon lorsqu'on cherche à les inférer à toute la population. Le seuil de signification alpha (0,05) est simplement obtenu en soustrayant du niveau de confiance (0,95) la valeur 1 : $\alpha_{0,05} = 1 - 0,95$. Aussi, lorsqu'un test d'hypothèse est significatif au seuil de 0,05 (5%), l'intervalle de confiance à 95% ne contient pas la **valeur de la relation nulle**. À l'inverse, lorsqu'un test n'est pas significatif, l'intervalle contient la **valeur de la relation nulle**.

NB : Le niveau de confiance (estimation) et le seuil de signification (test) donnent finalement la même information, mais sur une base complémentaire, le seuil de signification pouvant être vu comme un niveau de risque d'erreur 😊

2) Quels sont les liens qui existent entre l'analyse de tableaux bivariés et le test du chi-carré ?

D'une part, le test du chi-carré approfondit l'analyse tabulaire bivariée en aidant à établir dans quelle mesure une relation est significative au niveau de la population.

D'autre part, **avant ou après avoir établi la signification statistique d'une relation (chi-carré), il faut toujours en préciser la nature (tableau bivarié).**

NB : Le test du chi-carré ne précise pas la nature des différences. Par exemple, pour savoir si les femmes croient davantage au changement climatique que les hommes, il faut se référer à l'analyse tabulaire bivariée 😊

3) Un chercheur étudie la relation entre le temps passé à pratiquer du sport (variable indépendante) et l'indice de masse corporelle (variable dépendante). Il veut utiliser le test d'hypothèse du chi-carré. Précisez la nature et les valeurs possibles des variables (indépendante et dépendante) que le chercheur doit obtenir pour que le test en question soit approprié.

La variable indépendante, « temps de pratique de sport », doit être de nature **qualitative ou catégorielle** et ses valeurs possibles sont : **temps faible, temps moyen, temps élevé.**

La variable dépendante, « indice de masse corporelle », doit être de nature **qualitative ou catégorielle** et ses valeurs possibles sont : **IMC faible, IMC normal, IMC élevé.**

NB : À l'instar de l'analyse tabulaire bivariée, le test du chi-carré est approprié lorsque les variables indépendante et dépendante sont de nature catégorielle. Si les valeurs des variables sont de nature quantitative ou nombreuses, il faut les regrouper de telle sorte qu'elles se prêtent à l'analyse de tableaux bivariés 😊

- 4) Dans l'exemple de recherche suivant, strictement parlant, est-il justifié d'utiliser un test du chi-carré afin de tester s'il y a un lien d'association entre les deux variables ? Expliquez bien votre réponse.

Une étude de marché est menée en été 2011 dans le but de jauger l'intérêt de l'implantation d'un programme de maîtrise en santé mentale à la TÉLUQ. Un des enjeux est de savoir s'il y a un lien d'association entre le fait de recevoir une formation en santé mentale et l'intention de s'inscrire à ce nouveau programme. Parmi les personnes ciblées par le sondage web, 1667 participants (étudiants et professionnels de la santé) ont pu **volontairement** compléter le sondage. Le tableau ci-dessous présente les fréquences observées:

Tableau des fréquences observées

Intention de s'inscrire	Formation en santé mentale	
	Non, ma formation n'est pas liée à la santé mentale	Oui, ma formation est liée à la santé mentale
Oui	217	1067
Non, jamais	145	238

Il n'est pas justifié de recourir à un test du chi-carré pour au moins une raison : le caractère non probabiliste du mode d'échantillonnage choisi, puisque les répondants ont volontairement complété le sondage web. Or, nous savons qu'avec ce genre d'échantillonnage, tous les individus de la population recensée n'ont pas une chance non nulle indépendante et égale d'y faire partie. Aussi, il ne fournit ni l'assurance d'une forte représentativité de l'échantillon ainsi constitué ni la possibilité d'une estimation précise des biais en découlant inévitablement.

NB : Il peut arriver que le chercheur utilise le test du chi-carré dans le cas d'un échantillon non probabiliste. Dans ce cas, il doit être conscient des limites de l'inférence statistique et des implications sur les conclusions de sa recherche ☺

- 5) Est-il justifié d'utiliser un test du chi-carré afin de tester s'il y a un lien d'association entre la présence d'enfants et le statut civil chez 2173 étudiants sélectionnés de façon aléatoire ? Expliquez bien votre réponse.

Tableau des fréquences observées

Statut civil	Présence d'enfants	
	Oui	Non
Célibataire	37	1771
En couple	93	128
Séparé(e)/divorcé(e)	33	11

Tableau des fréquences théoriques

Statut civil	Présence d'enfants	
	Oui	Non
Célibataire	142,2	1665,8
En couple	17,4	303,6
Séparé(e)/divorcé(e)	3,5	40,5

En principe, il n'est pas justifié de recourir à un test du chi-carré pour au moins une raison liée au fait qu'une des fréquences théoriques est inférieure à 5 (**règle de Fischer**). En fait, les fréquences théoriques inférieures à 5 posent un sérieux problème au calcul du chi-carré. Elles ont tendance à diminuer le chi-carré, dont les valeurs possibles sont justement mathématiquement limitées.

Pour éviter ce problème, une des solutions consiste à fusionner des catégories moins fréquentes de façon à augmenter les fréquences théoriques. L'autre revient à exclure tout simplement les catégories présentant de faibles fréquences.

Toutefois, si on peut tolérer des fréquences théoriques inférieures à 5 pourvu qu'elles ne dépassent pas 20% des classes ou cellules (**règle de Cochran**), on ne peut en aucun cas accepter des fréquences théoriques nulles. À cet égard, puisque seule 1 cellule sur 6 (16,7%) comporte une fréquence théorique non nulle inférieure à 5, on peut tolérer cet inconvénient et effectuer le test du chi-carré puisque l'échantillon est aléatoire et de grande taille ($n=2173$).

NB : Pour utiliser le chi-carré, catégoriser les valeurs de telle sorte qu'il n'y ait pas de fréquences théoriques < 5 ; l'importance de la taille de l'échantillon aide en ce sens. À défaut, lorsqu'une fréquence théorique est < 5 , on peut utiliser le test exact de Fischer, lequel n'est pas couvert par ce cours 😊

6) On s'interroge sur la représentativité des femmes dans l'enseignement universitaire dans le domaine des sciences. Pour en savoir davantage, une étude institutionnelle porte sur la relation entre le genre et le rang académique des professeurs. Le sondage mené auprès de 65 professeurs d'une Faculté de Génie donne le tableau bivarié en fréquences ci-dessous :

Rang académique	Sexe		Total
	Femme	Homme	
Professeurs assistants	11	20	31
Professeurs agrégés	7	15	22
Professeurs titulaires	2	10	12
Total	20	45	65

a) Parmi les femmes, quel est le pourcentage de professeurs assistants? Parmi les femmes, quel est le pourcentage de professeurs titulaires?

Parmi les femmes, le pourcentage de professeurs assistants est de $((11/20)*100)$, soit 55%. Parmi les femmes, le pourcentage de professeurs titulaires est de $((2/20)*100)$, soit 10%.

b) Parmi les hommes, quel est le pourcentage de professeurs assistants? Parmi les hommes, quel est le pourcentage de professeurs titulaires?

Parmi les hommes, le pourcentage de professeurs assistants est de $((20/45)*100)$, soit 44,4%. Parmi les hommes, le pourcentage de professeurs titulaires est de $((10/45)*100)$, soit 22,2%.

c) En comparant les pourcentages ci-dessus, concluez sur la relation.

La différence entre les femmes et les hommes quant au rang de professeurs assistants est de 10,6 en points de pourcentages (55%-44,4%). De même, la différence entre les

femmes et les hommes quant au rang de professeurs titulaires est de 12,2% en points de pourcentages (22,2%-10%). Par conséquent, il semble qu'il y ait une relation faible entre le sexe et le rang académique dans la Faculté.

- d) Vous jugez qu'il est nécessaire de recourir au test d'hypothèse du chi-carré. Pour ce faire, formulez les hypothèses nulle et alternative.

H_0 : Le sexe et le rang académique des professeurs ne sont pas associés dans la population étudiée (absence de relation, de différences donc).

H_1 : Le sexe et le rang académique des professeurs sont associés dans la population étudiée (existence d'une relation, de différences donc).

- e) Calculez les fréquences théoriques ou attendues

Rang académique	Sexe		Total
	Femme	Homme	
Professeurs assistants	9,54 (31*20/65)	21,46 (31*45/65)	31
Professeurs agrégés	6,77 (22*20/65)	15,23 (22*45/65)	22
Professeurs titulaires	3,69 (12*20/65)	8,31 (12*45/65)	12
Total	20	45	65

- f) Calculez le chi-carré

F_0	F_a	$F_0 - F_a$	$(F_0 - F_a)^2$	$(F_0 - F_a)^2 / F_a$
11	9,54	1,46	2,14	0,22
7	6,77	0,23	0,05	0,01
2	3,69	-1,69	2,86	0,78
20	21,46	-1,46	2,14	0,10
15	15,23	-0,23	0,05	0,00
10	8,31	1,69	2,86	0,34
Total		0,00	$\chi^2 =$	1,46

- g) Calculez le nombre de degrés de liberté du chi-carré

Le nombre de degrés de liberté est de (3-1)(2-1), soit 2.

- h) Déterminez la valeur critique du chi-carré au seuil de signification de 0,05

Au seuil de signification de 0,05 et avec 2 degrés de liberté, la valeur critique du chi-carré est de 5,99.

- i) Prenez une décision quant au rejet ou l'acceptation de l'hypothèse nulle et concluez sur la signification statistique de la relation

La valeur calculée du chi-carré (1,46) étant plus petite que la valeur critique du chi-carré (5,99), nous acceptons l'hypothèse nulle. Autrement, l'écart n'est pas suffisamment important pour ne pas être dû au hasard de l'échantillonnage.

Finalement, il n'existe pas de relation statistique significative entre le sexe et le rang académique des professeurs. On est sûr au moins à 95% de l'inexistence de cette relation dans la population. Il y a donc 5% de chance de se tromper.

- j) Rédigez un texte d'environ 10 lignes dans lequel vous présentez vos résultats en précisant les différences observées entre hommes et femmes, le test du chi-carré comprenant la valeur du chi-carré, les dl, le seuil de signification (présentation des résultats= **interprétation statistique**).

L'analyse tabulaire bivariée indique que la différence entre les femmes et les hommes quant au rang de professeurs assistants est de 10,6 en points de pourcentages (55%-44,4%). De même, la différence entre les femmes et les hommes quant au rang de professeurs titulaires est de 12,2% en points de pourcentages (22,2%-10%). Il semble qu'il y ait une relation, si faible soit-elle, entre le sexe et le rang académique dans les données portant sur les 65 professeurs de la Faculté de Génie. Toutefois, les différences s'avèrent non significatives statistiquement au seuil de 0,05 ($\chi^2 = 1,46$, dl = 2). Par conséquent, on est sûr au moins à 95% que le sexe n'est pas associé au rang académique dans la population.

- k) Discutez les résultats de l'analyse en expliquant le pourquoi de la présence ou de l'absence d'une relation entre le sexe et le rang académique des professeurs (discussion des résultats = **interprétation sociologique/théorique**).

Si l'analyse tabulaire bivariée a montré une faible relation entre le sexe et le rang académique des professeurs dans n , le test du chi-carré a démontré que cette relation n'était pas significative au niveau de N . Cette **découverte négative** peut s'expliquer, en grande partie, par la faiblesse relative de la taille de l'échantillon n , lequel n'est constitué que de 65 cas. Il aurait fallu augmenter le nombre de cas sondés pour que la relation devienne probablement significative. Car, des études ont montré que les femmes sont encore victimes de discrimination lorsqu'il s'agit d'atteindre des postes de rang professoral et d'accéder à une carrière académique dans le domaine des sciences « exactes ». Par ailleurs, les sociologues ont expliqué comment la socialisation influençait le domaine de formation et l'orientation professionnelle des femmes vers l'univers des sciences sociales et humaines.

NB : Chaque fois que vos analyses vous conduisent à des découvertes négatives, vous devez expliquer cet écart entre les résultats escomptés et les résultats obtenus afin d'éclairer le terrain pour d'autres chercheurs. Par exemple, un chercheur pourrait reprendre la même étude en augmentant la taille de l'échantillon ou en effectuant une étude par recensement ☺

- 7) Une enquête menée en 2006 dans la zone épidémiologique du sud-ouest de Montréal (ZEPSOM) donne le tableau bivarié en % ci-dessous. La question de recherche est de savoir si le statut civil influe sur la prévalence de la détresse psychologique dans la population.

Tableau 1. Détresse psychologique selon le statut civil en % (n=239)

Détresse psychologique élevée	Statut civil	
	Célibataire/veuf /divorcé /séparé	Vivant avec un partenaire
Oui	43,5	30,6
Non	56,5	69,4
Total %	100,0	100,0
Nombre de cas (n)	131	108

Source : Enquête ZEPSOM, 2009

- a) Pour répondre à la question de recherche, commencez par analyser le tableau bivarié en % (interprétation statistique)

L'analyse tabulaire bivariée montre que 43,5% des répondants montréalais célibataires, veufs, divorcés ou séparés présentent une détresse psychologique élevée au cours des 12 derniers mois précédant l'enquête, comparativement à 30,6% chez les répondants vivant avec un partenaire. La différence en points de pourcentages étant d'environ 13, nous concluons à l'existence d'une relation d'association, si faible soit-elle, entre la détresse psychologique et le statut civil chez les 239 répondants constituant l'échantillon.

- b) Vous jugez nécessaire d'effectuer un test du chi-carré. Pour ce faire, formulez les hypothèses statistiques nulle et alternative

H_0 : La détresse psychologique n'est pas associée au statut civil dans la population du sud-ouest de Montréal (absence de relation, de différences donc).

H_1 : La détresse psychologique est associée au statut civil dans la population du sud-ouest de Montréal (existence d'une relation, de différences donc).

- c) Transformez les pourcentages du tableau bivarié en fréquences observées en montrant les calculs et en y intégrant les fréquences marginales

Pour transformer les % en fréquences absolues, je pose l'équation suivante :

Par exemple, pour la première cellule, que représente 43,5% de 131 cas ? La réponse est vite trouvée : je multiplie 43,5% par 131 (Excel) ou je multiplie 43,5 par 131 et je divise le dividende par 100 pour obtenir 57. De même, que représente 30,6% de 108 cas ? La réponse est vite trouvée : je multiplie 30,6% par 108 (Excel) ou je multiplie 30,6 par 108 et je divise le dividende par 100 pour obtenir 33. Ainsi de suite.... Voici le tableau des calculs et le tableau bivarié en fréquences.

Tableau de calcul des fréquences observées à partir des %

Détresse élevée	Statut civil	
	Célibataire/veuf /divorcé /séparé	Vivant avec un partenaire
Oui	$(43,5 \cdot 131) / 100$	$(30,6 \cdot 108) / 100$
Non	$(56,5 \cdot 131) / 100$	$(69,4 \cdot 108) / 100$
Nombre de cas (n)	131	108

Détresse psychologique selon le statut civil en fréquences observées (n=239)

Détresse élevée	Statut civil		Total
	Célibataire/veuf /divorcé /séparé	Vivant avec un partenaire	
Oui	57	33	90
Non	74	75	149
Total	131	108	239

Source : Enquête sud-ouest Montréal, 2009

d) Calculez les fréquences théoriques ou attendues en montrant les calculs.

Déresse élevée	Statut civil		Total
	Célibataire/veuf /divorcé /séparé	Vivant avec un partenaire	
Oui	$(90 \cdot 131) / 239 = 49,3$	$(90 \cdot 108) / 239 = 40,7$	90
Non	$(149 \cdot 131) / 239 = 81,7$	$(149 \cdot 108) / 239 = 67,3$	149
Total	131	108	239

e) Calculez le chi-carré et dites laquelle des combinaisons de catégories apporte une contribution plus importante à la valeur du chi-carré.

F ₀	F _a	F ₀ - F _a	(F ₀ - F _a) ²	(F ₀ - F _a) ² / F _a
57	49,3	7,7	59,3	1,2
74	81,7	-7,7	59,3	0,7
33	40,7	-7,7	59,3	1,5
75	67,3	7,7	59,3	0,9
		0,0	$\chi^2 =$	4,3

La combinaison « vivant avec un partenaire » et « oui : détresse élevée » apporte une contribution plus importante à la valeur du chi-carré, puisque cette contribution est symbolisée par **1,5** (soit l'écart au carré entre la fréquence observée 33 et la fréquence théorique 40,7, relativisé par la fréquence théorique 40,7).

f) Calculez le nombre de degrés de liberté du chi-carré et déterminer la valeur critique du chi-carré à un seuil déterminé.

$$DI = (2-1)(2-1) = 1$$

Avec 1 degré de liberté, la valeur critique du chi-carré au seuil de 0,05 est de 3,83.

g) Calculez le coefficient de contingence

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}}$$

$$C = \text{Racine } (4,3 / (4,3 + 239))$$

$$= \text{Racine } (0,018) = 0,13$$

h) Répondez à la question de recherche en prenant une décision quant au rejet ou l'acceptation de l'hypothèse nulle, en concluant sur la signification statistique et réelle de la relation.

La valeur calculée du chi-carré (4,3) étant plus grande que la valeur critique du chi-carré (3,83) avec 1 degré de liberté et au seuil de 0,05, je rejette l'hypothèse nulle d'une absence de relation. Par conséquent, on est sûr au moins à 95% que la prévalence de la détresse psychologique est significativement associée au statut civil dans la zone épidémiologique du sud-ouest de Montréal, quoique l'association statistique demeure faible (la valeur du coefficient de contingence étant de seulement 0,13).

8) Vous souhaitez vérifier l'hypothèse de recherche selon laquelle la croyance

concernant le changement climatique varie selon le sexe : les femmes croient davantage au réchauffement de la terre que les hommes. À partir d'un sondage aléatoire mené par Léger Marketing en hiver 2011 sur 1138 Canadiens, vous obtenez le tableau bivarié en fréquences ci-dessous:

		Gender		Total
		Male	Female	
Is there solid evidence that the average temperature on earth has been getting warmer over the past four decades?	No	94	78	172
	Yes	461	505	966
Total		555	583	1138

a) Spécifiez les variables indépendante et dépendante.

Le sexe est la variable indépendante et la croyance au changement climatique, la variable dépendante.

b) Construisez le tableau bivarié en pourcentages.

Tableau 1. Croyance au changement climatique selon le sexe en % (n=1138)

Croyance	Sexe	
	Homme	Femme
Non	16,9	13,4
Oui	83,1	86,6
Total %	100,0	100,0
Nombre de cas (n)	555	583

Source : Léger Marketing, Hiver 2011

c) Selon l'analyse tabulaire bivariée, les femmes sont-elles plus susceptibles de croire au réchauffement de la terre que les hommes? Justifiez-vous.

L'analyse tabulaire bivariée laisse entrevoir une relation nulle entre le sexe et la croyance au changement climatique. En fait, parmi les hommes, 83,1 % croient au changement climatique, tandis que chez les femmes, 86,6% croient à ce changement climatique. L'écart n'est donc pas important : il est seulement de 3,5 points de pourcentages. Mais, effectuons quand même le test pour vérifier.

d) Effectuez le test d'hypothèse du chi-carré au seuil de signification de 0,05 en respectant les différentes étapes.

1) Je formule les hypothèses nulle et alternative du test

H_0 : Le sexe et la croyance au changement climatique sont indépendants dans la population canadienne (absence de relation)

H_1 : Le sexe et la croyance au changement climatique sont dépendants dans la population canadienne (existence d'une relation)

2) Je calcule les fréquences théoriques ou attendues selon l'hypothèse nulle

		Gender		Total
		Male	Female	
Is there solid evidence that the average temperature on earth has been getting...?	No	83,88 (172*555/1138)	88,11 (172*583/1138)	172
	Yes	471,11 (966*555/1138)	494,88 (966*583/1138)	966
Total		555	583	1138

3) Je calcule le chi-carré

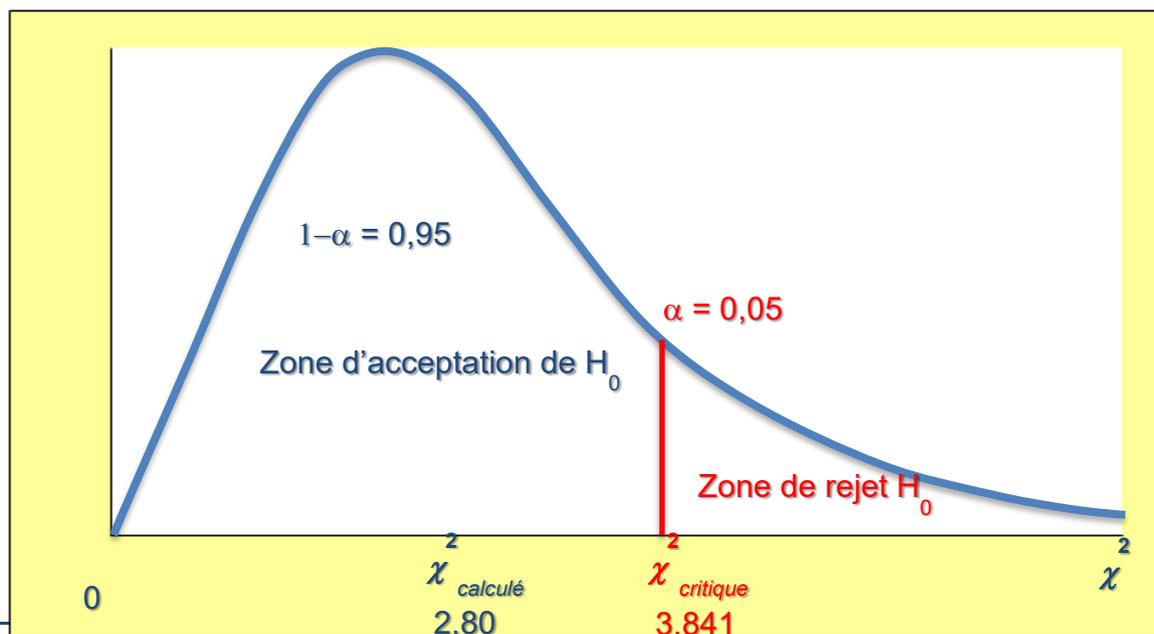
Fo	Fa	Fo - Fa	(Fo - Fa) ²	(Fo - Fa) ² / Fa
94	83,88	10,12	102,41	1,22
461	471,11	-10,11	102,21	0,22
78	88,11	-10,11	102,21	1,16
505	494,88	10,12	102,41	0,21
$\chi^2 =$				2,80

4) Je calcule le nombre de degrés de liberté du chi-carré

$$dl = (2-1)(2-1) = 1$$

5) Je fixe la règle de décision et représente graphiquement la décision

Je rejette H_0 , si la valeur calculée du chi-carré est supérieure à la valeur critique. La courbe ci-dessous réfère à la **distribution d'échantillonnage du chi-carré**. Elle se lit ainsi : *s'il n'y avait pas de relation entre le sexe et la croyance au changement climatique dans la population canadienne, 95% des chi-carrés des échantillons possibles se situeraient entre 0 et 3,841*. Or, le chi-carré calculé sur la base des données de l'échantillon se situe dans cette zone de nullité de la relation.



6) Je prends une décision et dégage une conclusion sur la relation

La valeur calculée du chi-carré (2,80) étant plus petite que la valeur critique du chi-carré (3,841) avec 1 degré de liberté et au seuil de 0,05, j'accepte l'hypothèse nulle. L'écart n'est pas suffisamment important pour qu'il ne soit pas dû au hasard. Par conséquent, on est sûr au moins à 95% qu'il n'y a pas de relation statistiquement significative entre le sexe et la croyance au changement climatique.

e) Le test est-il significatif au seuil de 0,10 (10% de risque d'erreur)?

La valeur calculée du chi-carré (2,80) étant plus grande que la valeur critique du chi-carré (2,706) avec 1 degré de liberté et au seuil de 0,10, je rejette l'hypothèse nulle pour accepter l'hypothèse alternative à l'effet qu'il y a une relation de dépendance entre le sexe et la croyance au changement climatique dans la population. Le test d'hypothèse est donc significatif au seuil de 0,10. On est sûr au moins à 90% que cette relation existe dans la population.

NB : On se rend compte que plus le seuil de signification est grand (moins exigeant), plus les chances que le test soit significatif augmentent. Le seuil doit donc être choisi en fonction de la gravité des implications de la décision. Par exemple en aéronautique ou en médecine, le seuil s'avère plus exigeant qu'en sciences sociales du fait de la sévérité des conséquences associées aux erreurs 😊

9) Je présente ainsi les résultats d'une analyse que j'ai menée (interprétation statistique):

L'analyse des données de la présente recherche (n= 2404) indique que les femmes sont 12,6% à consommer des somnifères, alors que les hommes sont 7,9% à s'adonner à cette consommation. La différence entre les deux groupes est de seulement 4,7%, mais elle s'avère significative (chi-carré = 13,99, dl = 1). On peut donc conclure à 99,9% que le sexe influence la consommation de somnifères chez les habitants âgés de 15 ans et plus de la zone épidémiologique du sud-ouest de Montréal.

a) Question de recherche

La consommation de somnifères varie-t-elle selon le sexe ?

b) Variables indépendante et dépendante, ainsi que leurs valeurs

Variable indépendante : Sexe (homme, femme)

Variable dépendante : Consommation de somnifères (oui, non)

c) Population étudiée

La population réfère à l'ensemble des habitants âgés de 15 ans et plus de la zone épidémiologique du sud-ouest de Montréal.

d) Taille de l'échantillon

L'échantillon est constitué de 2404 cas.

e) Exprimez en des termes sémantiques, c.-à-d. non techniques, la signification de la valeur du «chi-carré»

La valeur du chi-carré (13,99) signifie l'écart entre les fréquences observées d'une part et les fréquences attendues s'il n'y avait pas de relation entre le sexe et la

consommation de somnifères dans la population d'autre part.

- f) Exprimez en des termes sémantiques, c.-à-d. non techniques, la signification de la valeur du « dl »

Le nombre de degrés de liberté (1) signifie le nombre de rangées par colonnes du tableau bivarié à partir duquel on peut déduire les autres fréquences de cellule, connaissant les fréquences marginales de la distr. du sexe et de la consommation.

- g) Que vaut la valeur critique du chi-carré sur la table, au seuil de signification de 0,001 et avec 1 degré de liberté ? Que signifie cette valeur critique ?

La valeur critique du chi-carré vaut 10,83, soit la valeur minimale nécessaire au rejet de l'hypothèse nulle. Précisément, *si l'on obtenait tous les échantillons possibles (distribution d'échantillonnage), la valeur maximale du chi-carré équivaldrait à 10,83 s'il n'y avait pas de relation entre le sexe et la consommation de somnifères.*

- h) Exprimez en des termes sémantiques, c.-à-d. non techniques, la signification de « 99,9% »

La valeur 99,9% est le niveau de confiance du test d'hypothèse du chi-carré. Elle signifie qu'il y a au moins 99,9% de chances de ne pas nous tromper si nous concluons à l'existence d'une relation significative entre le sexe et la consommation.

- i) L'analyse tabulaire bivariée montre que la différence entre les hommes et les femmes est de seulement 4,7 en points de % (< 10). Pourtant cette différence s'avère statistiquement significative à un seuil très exigeant. Pourquoi ?

Une des explications est l'importance de la taille de l'échantillon (n=2404). Si la taille était petite, la différence entre les hommes et les femmes quant à la consommation de somnifères serait probablement non significative.

NB : Plus la taille de l'échantillon est grande, plus une relation tend à être statistiquement significative. Par conséquent, il faut déterminer la signification réelle de cette relation à l'aide d'une mesure d'association appropriée 😊