

Exercice (Khi-deux) :

Une étude porte sur le savoir-vivre (politesse, galanterie, maintien, etc.) dans la population américaine (fichier excel).

Une des questions posées aux répondants était:

Qu'est-ce qui importe plus socialement : permettre la liberté totale d'expression même si cela ouvre la porte aux mauvaises manières ou bien promouvoir les bonnes manières même si cela limite la liberté d'expression ?

Dans cet exemple, nous voulons tester :

H0 : indépendance entre les variables catégorielles SEXE et LIBERTE

→ L'indépendance ici signifierait que la proportion de répondants préférant la liberté d'expression serait la même pour les hommes que pour les femmes.

Solution (sous spss) :

Statistiques descriptives

D'après le tableau1, 935 participants ont donné une réponse valide à la question citée précédemment et à celle portant sur le fait qu'ils soient un homme ou une femme. Cela représente 92,6 % de l'échantillon (avec 75 observations manquantes). Le test sera donc basé sur ces observations.

Récapitulatif de traitement des observations

	Valide		Observations Manquant		Total	
	N	Pourcentage	N	Pourcentage	N	Pourcentage
sex Sex of respondent * freedman Which is more important to society:	935	92,6%	75	7,4%	1010	100,0%

Le deuxième tableau est un **tableau croisé**.

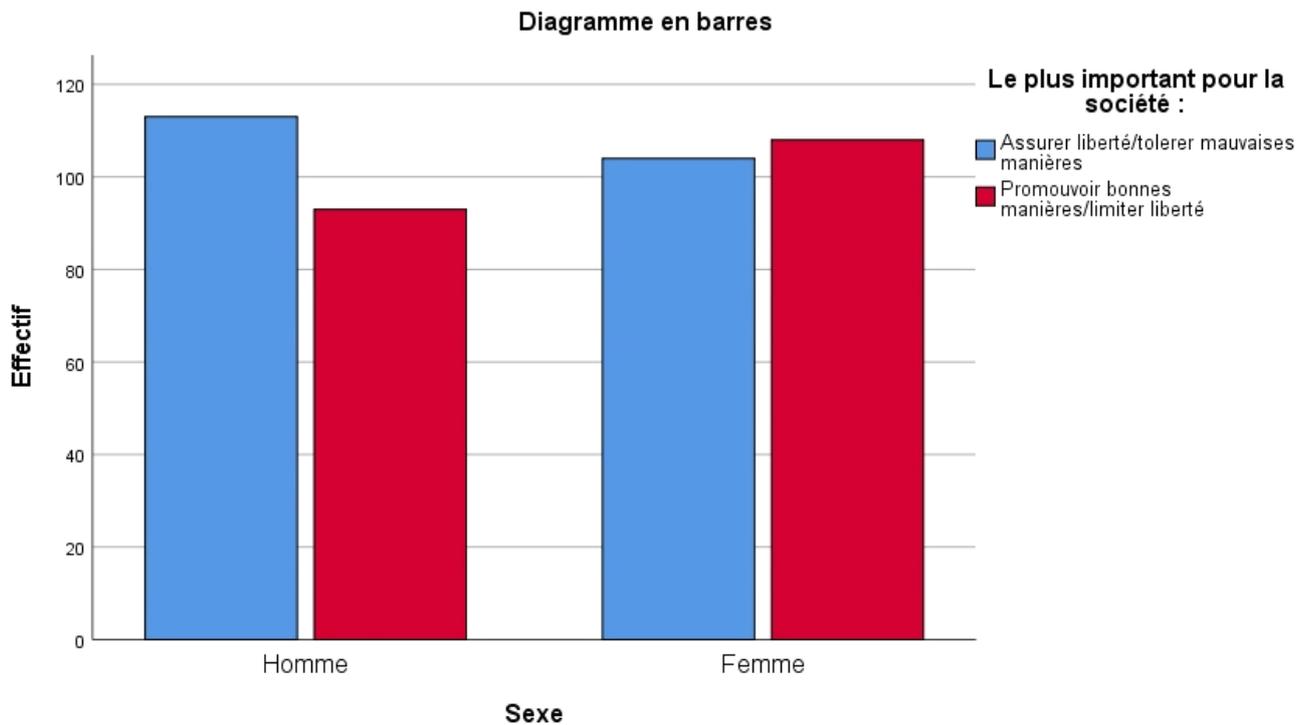
Tableau croisé sex Sex of respondent * freedman Which is more important to society:

sex Sex of respondent		freedman Which is more important to society:		Total	
		1 Allowing people freedom of expression/tolerate bad manners	2 Enforcing good manners/limit freedom of expression		
1 Male	Effectif	268	195	463	
	% dans sex Sex of respondent	57,9%	42,1%	100,0%	
	2 Female	Effectif	232	240	472
		% dans sex Sex of respondent	49,2%	50,8%	100,0%
Total	Effectif	500	435	935	
	% dans sex Sex of respondent	53,5%	46,5%	100,0%	

On remarque que **58 % des hommes** sont en faveur de **l'assurance de la liberté d'expression** contre **49 % pour les femmes**.

À partir de ce résultat, peut-on affirmer que les hommes et les femmes ont des attitudes différentes en regard de la liberté d'expression ?

Dans cet échantillon, les hommes supportent davantage la première option. Nous pouvons d'ailleurs observer les différences entre les hommes et les femmes dans le graphique à barres ci-dessous.



Nous voulons donc tester l'hypothèse nulle (si la variable LIBERTE est indépendante de la variable SEXE).

Nous devons donc calculer la **statistique du Chi-2** à partir de la différence entre **les occurrences observées et attendues**.

Puisque SPSS peut le calculer pour nous, nous avons refait la procédure en demandant ces trois informations :

- les occurrences observées et attendues et la différence entre les deux (résiduel non standardisé).

Tableau croisé sex Sex of respondent * freedman Which is more important to society:

		freedman Which is more important to society:		Total	
		1 Allowing people freedom of expression/tolerate bad manners	2 Enforcing good manners/limit freedom of expression		
sex Sex of respondent	1 Male	Effectif	268	195	463
		Effectif théorique	247,6	215,4	463,0
	% dans sex Sex of respondent	57,9%	42,1%	100,0%	
	Résidu	20,4	-20,4		
2 Female	2 Female	Effectif	232	240	472
		Effectif théorique	252,4	219,6	472,0
	% dans sex Sex of respondent	49,2%	50,8%	100,0%	
	Résidu	-20,4	20,4		
Total	Total	Effectif	500	435	935
		Effectif théorique	500,0	435,0	935,0
		% dans sex Sex of respondent	53,5%	46,5%	100,0%

Nous remarquons que l'effectif **théorique (ou attendu)** se distingue de l'effectif **observé** dans l'échantillon. **S'il n'y avait pas de relation entre les deux variables, l'effectif et l'effectif théorique seraient identiques.**

La différence est-elle significative ? Nous pouvons évaluer le tout à partir de la statistique Chi-2.

Pour ce faire, nous devons faire la somme de chaque résidu (occurrence observée – occurrence attendue) mis au carré divisé par l'occurrence attendue. Voici le calcul :

$$\frac{(268 - 247,6)^2}{247,6} + \frac{(195 - 215,4)^2}{215,4} + \frac{(232 - 252,5)^2}{252,5} + \frac{(240 - 219,6)^2}{219,6} = 7,16$$

On remarque que cette formule calcule l'écart ou la distance entre l'hypothèse nulle (occurrences attendues) et la situation réelle (occurrences observées). La valeur de Chi-2 est une quantification de cet écart. Plus la valeur sera élevée, plus l'écart entre l'hypothèse nulle et la situation réelle sera grand.

Nous devons maintenant positionner ce résultat dans une distribution de Chi-2 ajustée en fonction du degré de liberté (nombre de rangées – 1) * (nombre de colonnes – 1). Dans ce cas-ci, le degré de liberté est 1.

La comparaison avec la distribution Chi-2 est effectuée par SPSS.

Résultat du Chi-2

Ce tableau de résultat apparaît tout de suite sous le tableau croisé.

Tests du khi-carré

	Valeur	ddl	Signification asymptotique (bilatérale)	Sig. exacte (bilatérale)	Sig. exacte (unilatérale)
khi-carré de Pearson	7,161 ^a	1	,007		
Correction pour continuité ^b	6,815	1	,009		
Rapport de vraisemblance	7,171	1	,007		
Test exact de Fisher				,009	,005
Association linéaire par linéaire	7,154	1	,007		
N d'observations valides	935				

a. 0 cellules (0,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 215,41.

b. Calculée uniquement pour une table 2x2

On remarque que la valeur de Chi-2 calculée par SPSS est identique à celle que nous avons calculée à la main.

On observe aussi que le degré de signification est très bas, ce qui indique que les différences entre les occurrences observées et attendues sont significatives, ce qui veut dire que l'on retrouverait ces différences 7 fois sur 1000 si l'hypothèse nulle était vraie.

On doit donc rejeter l'hypothèse nulle selon laquelle les femmes et les hommes répondent pareillement à la question. Il appert que les femmes préfèrent les bonnes manières significativement plus souvent que les hommes.

La dernière rangée indique le nombre d'observations valides qui ont été utilisées pour calculer la valeur du Chi-2.

Dans le bas du tableau, on voit que SPSS nous informe si nous respectons la prémisse concernant le nombre minimal d'occurrences attendues par cellules. Nous voyons que dans notre cas, nous respectons la prémisse, puisque la plus petite occurrence attendue est de 215,41.

La taille d'effet

Les mesures d'associations (**Phi et V de Cramer**) nous indiquent **la force de la relation entre le sexe des participants et leur réponse à la question portant sur les bonnes manières et la liberté d'expression.**

Mesures symétriques

Nominal par Nominal	Phi	,082	,013
	V de Cramer	,082	,013
N d'observations valides		910	

La statistique Phi (pertinente dans le cas d'un tableau 2×2) a une valeur de 0,082 sur une valeur maximale possible de 1. Cela représente une petite association entre le sexe des participants et le fait qu'ils privilégient les bonnes manières plutôt que la liberté d'expression (si vous pensez en termes de coefficient de corrélation, cette valeur représente un effet de petite taille).

La valeur du test est très significative ($p < 0,01$), ce qui indique que les chances d'obtenir une valeur de 0,082 par hasard est très faible.

Nous confirmons donc le résultat du Chi-2 : la relation entre le sexe et la variable LIBERTE est statistiquement significative, mais de faible magnitude.